Introducción a la teoría de gramáticas. Lenguajes y autómatas

Jhenny Castillo Tapia jhennyct@yahoo.es

INTRODUCCIÓN

El universo de los sistemas automatizados, digital, virtual. Cada día absorbe más terreno y forma parte ya de nosotros. Por ejemplo en el interior de nuestros hogares, encontramos lavadoras, hornos microondas equipos de sonido, video, sistemas de calefacción/aire acondicionado etc. En el lugar menos pensado, la cocina o el cuarto de lavado, usted posee tecnología de punta, sistemas automatizados obras de ingeniería, pero que, por la familiaridad o uso frecuente parece algo normal y parte de nosotros incluso nuestros niños interactúan cada día en Video Juegos, TV, DVD, Play Stations y otros con la mayor familiaridad, sin imaginarse la complejidad de los sistemas que nos rodean en el mundo real.

Hasta hace unos años atrás se empleaba la frase "**Lento pero seguro**" con el objetivo de buscar la excelencia pensando que los trabajos realizados a la rápida no tenían buena calidad, empero con los sistemas automatizados surge una nueva premisa "**Rápido y muy seguro**". El universo tecnológico cada día busca satisfacer las necesidades del ser humano, mejorar y optimizar la forma de hacer las cosas, más rápido y mejor.

En lo que atañe a nuestro tema de estudio no es otra cosa que la comprensión y representación de estos dos universos paralelos el real y el virtual, el tema de estudio abarca más allá de las cosas tangibles o visibles, la expresión corporal, el lenguaje la forma de decir y comprender el significado y la semántica de las palabras generando lenguajes de comunicación. Por ejemplo:

Sean las siguientes expresiones:

Ven	Se refiere a la acción de venir.
Ven!	Es una orden, una expresión imperativa.
Ven?	Es una consulta, o pregunta que implica la voluntad del otro de realizarlo o no.

En el universo del lenguaje existen innumerables expresiones y formas de decir las cosas por lo que se hace necesaria la estandarización o formalización de un lenguaje, de manera que pueda adquirir un significado real para nosotros.

Si nos preguntamos qué significa lenguaje formal, probablemente lleguemos a analizar primeramente el término lenguaje, que en palabras sencillas, viene a ser todo tipo de expresión que permite al ser humano comunicarse con sus semejantes, con el cual podemos expresarnos y dar a conocer nuestras opiniones, es sin duda importante aclarar, que para poder entendernos entre personas, debemos hablar un mismo idioma.

Caso contrario podríamos caer en ambigüedades que ocasionaría malos entendidos, por ejemplo, el movimiento de girar la cabeza de izquierda a derecha o viceversa significa para nosotros en este lado del mundo "No", para otros podría significar "Si", otro ejemplo, es la forma de lectura/escritura lo normal para nosotros es realizarlo de izquierda a derecha, esto parece obvio, sin embargo no lo es, si consideramos que hay millones de personas que leen de derecha a izquierda, de arriba abajo o incluso en diagonal. Si analizamos la forma de escritura en Arabia o China tendríamos que analizar la forma de lectura, antes de realizarla, de lo contrario podríamos darle un significado muy diferente, como, sucedió en una publicidad de analgésicos (pastillas) para calmar el dolor de cabeza, se trabajó con las siguientes imágenes, bajo la premisa: "una imagen habla más que mil palabras"

Para lo que emplearon un cuadro con tres fotografías dispuestas de manera horizontal, en la primera aparece la imagen de una persona triste, en la segunda una persona tomando la pastilla, y en la tercera una persona feliz, que si las leemos en este lado del mundo, el significado sería una persona enferma toma la pastilla y se sana, parece normal que lo entendamos de esta manera, sin embargo no se imaginaron los resultados devastadores en la comercialización de este producto y el fracaso total, puesto que en Arabia lo normal es leer de derecha a izquierda, por lo tanto, ellos leyeron en la primera imagen una persona feliz, luego toma la pastilla, y se pone triste o se enferma. Este problema se dio, debido a que no se tomaron las previsiones concernientes al contexto donde se realizaría la campaña publicitaria.

Por lo que intentar llegar a formalizar un lenguaje nos ayudará a su mejor comprensión, consigna o instrucción a resolver/.

Uno de los pioneros sobre lenguajes y gramáticas fue Avram Noam Chomsky, sin duda la figura más destacada de la lingüística moderna, tanto por sus fundamentos matemáticos como por sus teorías sobre el origen y la naturaleza de los lenguajes naturales.

En el campo de la informática, el concepto de Gramática Formal adquirió gran importancia para el desarrollo de lenguajes de programación, consiguientemente el desarrollo de autómatas y maquinas de Turing cobró vida en las últimas décadas, fortaleciendo el vínculo entre Electrónica e Informática, creando máquinas cada vez mas sofisticadas y menos complicadas para el usuario final.

El propósito de este material está dirigido a introducir a los estudiantes universitarios de las ramas de la Informática, en el fascinante mundo de los lenguajes y la lógica implícita en las máquinas del siglo XXI. Proporcionando una guía práctica con ejercicios resueltos que pretenden fortalecer el conocimiento de la teoría de gramáticas y lenguajes formales, en el entendido que cada solución propuesta en este material, no representa la única solución, existiendo muchas maneras de resolver el mismo ejercicio.

CAPÍTULO I
CONCEPTOS Y DEFINICIONES

Veamos algunos conceptos que nos permitirán conceptualizar la gramática

SÍMBOLO

Es una entidad abstracta, que no se va a definir. Normalmente los símbolos son letras (a,b,c,…z), dígitos (0,1,2…9) y otros caracteres (+,*,/,-,?...).

Un símbolo también puede estar formado por varias letras o caracteres, como las palabras reservadas de un lenguaje de programación son símbolos de dicho lenguaje. Ejemplo:

- a,b,c,#,+,-,*, then, begin, end, else, …

VOCABULARIO O ALFABETO

Un vocabulario o alfabeto es un conjunto finito de símbolos, no vacío. Para definir que un símbolo *a* pertenece a un alfabeto *V*, se utiliza la siguiente notación $a \in V$.

Los alfabetos se definen por enumeración de los símbolos que contienen, podemos ver los siguientes ejemplos:

- V1={A,B,C,D,E,F,…..,X,Y,Z}
- V2={a,b,c,d,0,1,2,3,4,*,#,+}
- V3={0,1}
- V4={if, then, begin, end, else, a,b,;,=,>}
- También se pueden definir las tablas ASCII y EBCDIC como los alfabetos de distintos ordenadores.

CADENA

Una cadena es una secuencia finita de símbolos de un determinado alfabeto.
Ejm. Tomando en cuenta los alfabetos o vocabularios definidos anteriormente, podemos decir que:
abcb es una cadena del alfabeto V2
a+2*b es una cadena del alfabeto V2
000111 es una cadena del alfabeto V3
If a>b then b=a; es una cadena del alfabeto V4

LONGITUD DE CADENA

La longitud de una cadena consiste en el número de símbolos pertenecientes a la cadena. Ejm. Tomando en cuenta los ejemplos de cadena podemos decir que:

- |abcb| es de longitud 4
- |a + 2*b| es de longitud 5
- |000111| es de longitud 6
- |if a>b then a=b;| es de longitud 9

CADENA VACÍA

Se denomina cadena vacía, que no tiene símbolos y se denota con λ, por lo que su longitud es :
$| \lambda | \rightarrow 0$

CONCATENACIÓN DE CADENAS

Sean A y B dos cadenas cualesquiera, se denomina concatenación de A y B a una nueva cadena AB constituida por los símbolos de la cadena A seguidos por los de la cadena B.
El elemento neutro de la concatenación es λ:
$A \lambda = \lambda A = A$

UNIVERSO DEL DISCURSO

El conjunto de todas las cadenas que se pueden formar con los símbolos de un alfabeto, se denomina *universo del discurso* V y se representa por W(V). Evidentemente W(V) es un conjunto infinito. La cadena vacía pertenece a W(V).Ejm:
Sea un alfabeto con una sola letra V={a}, entonces el universo del discurso es:
W(V) = {λ, a, aa, aaa, aaaa, ….}

que contiene infinitas cadenas.

GRAMÁTICA

Veamos algunos conceptos que nos ayuden a formular el concepto de gramática:
(Del lat. grammatĭca, y este del gr. γραμματική). f. Ciencia que estudia los elementos de una lengua y sus combinaciones. Arte de hablar y escribir correctamente una lengua. Estudio de una lengua regido por el principio de que todos sus elementos mantienen entre sí relaciones sistemáticas. La que trata de formular una serie de reglas capaces de generar o producir todas las oraciones posibles y aceptables de un *idioma* o lenguaje
Microsoft® Encarta® 2007. © 1993-2006 Microsoft Corporation. Reservados todos los derechos.
Una definición un tanto técnica: "La gramática es un ente formal para especificar, de una manera finita, el conjunto de cadenas de símbolos que constituyen un lenguaje". La gramática genera o describe un lenguaje.

AUTÓMATA

(Del latin. *automăta*, t. f. de *-tus,* y este del gr. αὐτόματος, espontáneo). m. Instrumento o aparato que encierra dentro de sí el mecanismo que le imprime determinados movimientos o respuestas. Máquina que imita la figura y los movimientos de un ser animado. Microsoft® Encarta® 2007. © 1993-2006 Microsoft Corporation. Reservados todos los derechos.
En el caso de los *Procesadores de Lenguaje* un autómata es una construcción lógica que recibe como entrada una cadena de símbolos y produce una salida indicando si dicha cadena pertenece o no a un determinado lenguaje.

LENGUAJE

Conjunto de sonidos articulados con que el hombre manifiesta lo que piensa o siente. Sistema de comunicación verbal. Manera de expresarse. Conjunto de señales que dan a entender algo. El lenguaje de los ojos, el de las flores. En Informática Conjunto de signos y reglas que permite la comunicación con un ordenador. Microsoft® Encarta® 2007. © 1993-2006 Microsoft Corporation. Reservados todos los derechos.
Podemos expresarlo de manera más sencilla como un conjunto de palabras ó cadenas de símbolos (palabras, oraciones, textos o frases) de un determinado alfabeto.

LENGUAJE VACÍO

Existe un lenguaje denominado *lenguaje vacío*, que es un conjunto vacío y que se denota por {Ø}. El lenguaje vacío no debe confundirse con un lenguaje que contenga una sola cadena, y que ésta sea la

cadena vacía, es decir {λ}, ya que el número de elementos (cardinalidad) de estos dos conjuntos es diferente.
Cardinal ({ Ø }) = 0
Cardinal ({ λ }) = 1

PALÍNDROMOS

Cadenas que se leen igual hacia delante, que hacia atrás. Por ejemplo, ORURO

CAPÍTULO II
LENGUAJES Y GRAMÁTICAS
LENGUAJE

Se denomina lenguaje a un conjunto de palabras de un determinado alfabeto.
También un lenguaje es un conjunto de cadenas de símbolos (palabras, oraciones, textos o frases).
Un lenguaje está compuesto por **Sintaxis:** (gramática), que define las secuencias de símbolos que forman cadenas válidas de un lenguaje. Y por **Semántica**, que es el significado de las cadenas que componen un lenguaje.

Ejemplo 1:

Sintaxis: A
Semántica: es un número natural.
Diferente sintaxis en diferentes lenguajes:
A: natural
A: es un número que pertenece al conjunto de |N={1,2,3..N}

Ejemplo 2:

Sintaxis:

```
if a=b then write(a, ‘ es igual a ’, b )
        else write(a, ‘ es distinto a ’, b )
```

Semántica:

Si se cumple la condición entonces se muestra un mensaje que ambos números son iguales.
Caso contrario, se escribe los número son distintos.

Gramática

Chomsky la define como: “Descripción formalizada de las oraciones de un lenguaje. Una gramática **genera** o **describe** un lenguaje.”

DEFINICIÓN FORMAL DE GRAMÁTICA

Una gramática es una cuádrupla:

G = (VT, VN, S, P)

donde:

VT= {conjunto finito de símbolos terminales}
VN={conjunto finito de símbolos no terminales}
S es el *símbolo inicial* y pertenece a VN
P= {conjunto de producciones o de reglas de derivación}

Todas las cadenas del lenguaje definidas por la gramática están formadas con símbolos del *vocabulario terminal* VT. El vocabulario terminal se define por enumeración de los símbolos terminales.
El *vocabulario no terminal* VN es el conjunto de símbolos introducidos como elementos auxiliares para la definición de la gramática, y que no figuran en las sentencias del lenguaje.
La intersección entre el vocabulario terminal y no terminal es el conjunto vacío:

$$\{VN\} \cap \{VT\} = \{\emptyset\}$$

La unión entre el vocabulario terminal y no terminal es el *vocabulario*.

$$\{VN\} \cup \{VT\} = \{V\}$$

En ocasiones es importante distinguir si un determinado vocabulario incluye o no la cadena vacía, indicándose respectivamente con superíndice + o superíndice *, tal como se muestra a continuación:

$$V+ = V - \{\lambda\}$$
$$V^* = V + \{\lambda\}$$

El *símbolo inicial* S es un símbolo no terminal a partir del cual se aplican las reglas de la gramática para obtener las distintas cadenas del lenguaje.

Las *producciones* P son las reglas que se aplican desde el símbolo inicial para obtener las cadenas del lenguaje. El conjunto de producciones P se define por medio de la enumeración de las distintas producciones, en forma de reglas o por medio de un metalenguaje.

Ej 1: Sea la gramática: G=(VT,VN,S,P) donde VT={a,b}, VN={S} y el conjunto de producciones es:

$S \rightarrow ab$
$S \rightarrow aSb$

Las cadenas de esta gramática están dadas por: ab, aabb, aaabbb, $aa\ldots a_n bb\ldots b_n$

Ej 2. Sea la gramática: G=({a,b,c,d}, {S,A,B},S,P) donde P son las producciones:

$S \rightarrow ASB$
$A \rightarrow b$
$aaA \rightarrow aaBB$
$S \rightarrow d$
$A \rightarrow aA$
$B \rightarrow dcd$

Las cadenas de esta gramática son: bddcd, abddcd, aabddcd, bbaddcddcddcd……

Ej 3: Sea la gramática: G=(VN, VT,S,P) donde:

VN={<número>, <dígito>}
VT={0,1,2,3,4,5,6,7,8,9}
S= <número>

Las reglas de producción P son:

<número>::=<dígito><número>
<número>::=<dígito>
<dígito>::=0| 1| 2| 3| 4| 5| 6| 7| 8| 9

DEFINICIÓN FORMAL DE LENGUAJE

El lenguaje L (G) generado por una gramática G es el conjunto de todas las sentencias que puede generar G. Es decir expresado formalmente:

$$L(G) = \{\eta \in VT^*/S \rightarrow \eta\}$$

Una sentencia pertenece a L (G) si:

- η son símbolos terminales (VT)
- La sentencia puede derivarse del símbolo inicial S aplicando las reglas de producción de la gramática

PROPIEDAD

Dos gramáticas son equivalentes si ambas generan el mismo lenguaje.

G1 y G2 son equivalentes si L(G1) = L(G2)

EJEMPLO 1:

Sea la gramática definida por
G1= ({S}, {0,1},S,P)
donde
P={(S→ 000S111), (0S1 →01)} .
Determinar el lenguaje que genera.

SOLUCIÓN

La única forma de generar sentencias es aplicando cualquier número de veces la primera producción y terminado con la aplicación de la segunda, así se obtiene el lenguaje.

$$S \rightarrow 000S111 \rightarrow 000000S111111 \rightarrow \ldots\ldots \rightarrow 0^{(3n-1)}0S11^{(3n-1)} \rightarrow 0^{(3n)}1^{(3n)}$$

Por consiguiente el lenguaje que genera esta gramática es el conjunto infinito de instrucciones que se indica a continuación:

$$L(G2)=\{0^{(3n)}1^{(3n)}/n>=1\}$$

Si la 2ª producción de la gramática del ejemplo 1 fuese S→01 el lenguaje sería:

SOLUCIÓN
$L(G2)=\{0^{(3n+1)}1^{(3n+1)}/\ n>=0\}$

EJERCICIOS RESUELTOS

1. Sea la gramática definida por

$G_1= (\{S\},\{a,b\},S,P)$

Donde:

$P=\{(S\rightarrow aSb), (S\rightarrow ab)\}$

Determinar el lenguaje que genera.

SOLUCIÓN

$L(G_1)=\{a^nb^n/\ n>=1\}$

2. Dadas las siguientes palabras del lenguaje determinar las reglas de producción que las genera.

yxxyx
x
xyyx
$(xyy)^nx$
$(yxxy)^nx$

SOLUCIÓN

$G_2= (\{S,A\}, \{x,y\},S,P)$

donde $P=\{(S\rightarrow xyAS),(xyA \rightarrow yxxy)(S\rightarrow x),(A\rightarrow y)\}$.

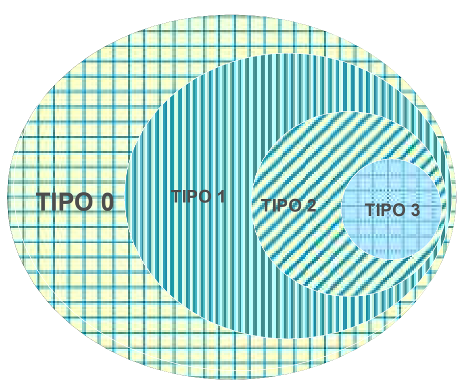
$L(G_2)$={cadenas que contienen xyy y yxxy intercambiándose y reproduciéndose cualquier número de veces, y terminado siempre con el símbolo x}

CAPÍTULO III
JERARQUÍA DE LAS GRAMÁTICAS

Para una mejor comprensión las gramáticas han sido clasificadas de acuerdo a particularidades y restricciones propias, una de ellas y la más acertada es la formulada por Avram Noam Chomsky, quien clasificó las gramáticas de acuerdo a cuatro tipos, dando origen a la **Jerarquía de Chomsky** en función de la forma de reglas de derivación o producción.

Las gramáticas no restringidas	Tipo 0
Sensibles al contexto	Tipo 1
Independientes del contexto	Tipo 2
Gramáticas regulares	Tipo 3

La clasificación comienza con un tipo de gramáticas que pretende ser universal, aplicando restricciones a sus reglas de derivación, se van obteniendo los otros tres tipos de gramáticas. Esta clasificación es jerárquica, es decir cada tipo de gramática incluye a todos los tipos siguientes.
Los lenguajes que resultan de dichas gramáticas también se identifican con lenguajes de tipo cero, uno, dos y tres.
A esta jerarquía de lenguaje se le conoce como la jerarquía de chomsky.



GRAMÁTICAS TIPO 0
También llamadas *gramáticas no restringidas con estructura de frase.* Se caracterizan por:

- En la parte izquierda tiene que haber al menos un símbolo no terminal.
- Respecto a sus partes derechas de producciones no hay ningún tipo de restricción.
- Las reglas de derivación son de la forma:

$$\beta \to \alpha$$

- Siendo $\alpha \in (VN \cup VT)+$ y $\in \beta$ $(VN \cup VT)$, es decir la única restricción es que no puede haber reglas de la forma

$$\lambda \to \beta$$

  donde λ es la cadena vacía.
- Ejemplos de estas gramáticas son todos los ejercicios que hemos visto hasta ahora.

EJEMPLO

- Sea la gramática definida por: G= ({S}, {0,1},S,P)
  donde
  P={(S→ 000S111), (0S1 →01)} .
  Determinar el lenguaje que genera.
  SOLUCIÓN
  $L(G)=\{0^{(3n+1)}1^{(3n+1)}/ n>=0\}$

GRAMÁTICAS TIPO 1
También llamadas *gramáticas sensibles al contexto.* Es decir que es importante tomar en cuenta la ubicación de los símbolos no terminales en la regla de derivación (que preceden y suceden a cada símbolo Terminal, deben mantener su ubicación en el lado derecho de la regla de producción tal como aparece en la parte izquierda de la regla de producción).
En este tipo de gramática sus reglas de producción son de la forma:

$$\alpha \mathbf{A} \beta \to \beta\gamma\alpha$$

Siendo $A \in VN$
$\alpha, \beta \in (VN \cup VT)^*$ y
$\gamma \in (VN \cup VT)^*$

Estas gramáticas se llaman sensibles al contexto, pues se puede reemplazar A por γ siempre que estén en el contexto $\alpha \ldots. \beta$.
EJEMPLOS

- La gramática G=({S,A,B}, {a,b}, S, P) cuyas producciones P se muestran a continuación:
  S → aB
  B → bAb
  bA→bcD
  cD →cAA
  A →aaA
  A →b

PROPIEDADES DE LAS GRAMÁTICAS TIPO 1

- **Propiedad de no decrecimiento**.- Las cadenas que se obtienen en cualquier derivación de una gramática de tipo 1 son de longitud no decreciente, es decir:

$$\alpha \to \beta \Rightarrow |\beta| \geq |\alpha|$$

  Y se puede enunciar como la longitud de la parte derecha de la producción es mayor o igual a la parte izquierda. Es decir no tiene reglas *compresoras*.
  Se puede demostrar de la siguiente manera:

$$\alpha A\ \beta \to \beta\gamma\alpha$$

Siendo $\in\gamma$ $(VN \cup VT)+$, es decir γ nunca puede ser la cadena vacía, lo que implica que $|\gamma| \geq 1$ y como |A| como mínimo vale 1, queda demostrada la propiedad:

$$|\alpha A\beta| <= |\alpha\gamma\beta|$$

- **Propiedad de sensibilidad al contexto**

En los lenguajes generados por estas gramáticas el significado de las "palabras" depende de su posición en la frase.
A los símbolos α y β es a lo que se llama **contexto**. Es decir, **A** sólo puede transformarse en γ si va precedido de α y seguido de β.

***Ejercicio** 1.*

- Dado el siguiente lenguaje elabora las reglas de producción
  $L(G) = \{ 0^n1^n / n \geq 1\}$

Solución

$G = \{ \{S, A\}, \{0, 1\}, S, P \}$
Reglas de producción:
$S \rightarrow 0SA$ / 01
$1A \rightarrow 11$

Ejercicio 2

- *Construye una gramática tipo 1 que genere el lenguaje*

$L = \{a(bc)^n / n >= 1\}$

Solución

S→ aB
B → bc B / bc

GRAMÁTICAS TIPO 2

- También se denominan *gramáticas de contexto libre o libres de contexto.*
- Sus reglas de producción tan sólo admiten tener un símbolo no terminal en su parte izquierda, es decir son de la forma:

$$\mathbf{A} \rightarrow \alpha$$

Siendo A $\in$ VN y $\in \alpha$ (VN $\cup$ VT) $^+$

Si cada regla se representa como un par ordenado (A, α), el conjunto P es un subconjunto del conjunto producto cartesiano VN x ({VN $\cup$ VT}) +, es decir:

PN{$\subset$ x ({VN} $\cup$ {VT})+}

La denominación contexto libre se debe a que se puede cambiar A por α, independientemente del contexto en que aparezca A.

EJEMPLOS

Ejemplo 1

- La gramática G=({S,A,B}, {a,b}, S, P) cuyas producciones P se muestran a continuación es de tipo 2:

S $\rightarrow$ aB
S $\rightarrow$ bA
A$\rightarrow$a
A $\rightarrow$ aS
A $\rightarrow$bAA
B $\rightarrow$b
B $\rightarrow$bS
B $\rightarrow$aBB

Ejemplo 2

- La gramática G=({a,b}, {A,S}, S,P) donde P son las producciones que se muestran a continuación es de tipo 2.
  S $\rightarrow$aS
  S $\rightarrow$aA
  A $\rightarrow$bA
  A $\rightarrow$b

Ejercicio 1.

- Dado el siguiente lenguaje $L(G) = \{ 0^n1^n / n \geq 1\}$

Con las siguientes reglas de producción tipo 1. Convertir a tipo 2

- *Reglas de producción:*
  $S \rightarrow 0SA$ / 01,
  $1A \rightarrow 11$

Solución

S → 0S1 / 01

Ejercicio 2

- *Construye una gramática tipo 2 que genere el lenguaje:*

  L = {a(bc)^n / n > = 1}

Solución

S→ aB
B → bc B / bc

GRAMÁTICAS TIPO 3

- También denominadas <u>*regulares o gramáticas regulares a la derecha*</u> comienzan sus reglas de producción por un símbolo terminal que puede ser seguido o no por un símbolo no terminal, es decir son de la forma:

A → aB
A → a

Donde A, B ∈ VN y α ∈ VT

EJEMPLO

Sea la gramática:
G=({a,b}, {A,S}, S, P)
donde P son las producciones formuladas por:

S → aS
S → aA
A → bA
A → b

Ejercicio 1

Construye una gramática tipo 2 que genere el lenguaje:
L = {a(bc)^n / n > = 1}

Solución:

S → aB/a
B → bC/bc
C→ cB

CAPÍTULO IV

EJERCICIOS RESUELTOS

1. Construye una gramática tipo0 que genere las cadenas V que no contengan la secuencia "abc"*

Solución:

S → aB/bB/cB/dB
B → bB/dB/λ
B→ aC/dC/λ
C→ dD/aC/aD/λ
D→ cD/dD/aD/λ
B→ bbC

2. Construye una gramática tipo 0 para el siguiente lenguaje:
L(G)={a^{n}b^m / n>=4,m>=3}

Solución:

S → aaaa/A
A → aA/b/bb/bbb/λ

3. Construye una gramática para el vocabulario V= {a,b,c,d} donde todas las cadenas generadas contengan una única a.

Solución:

S → aB/Ba
S → bC/cC
B→ bB/cB/λ
C→ bC/cC/a/Cb/Cc

4. L = {xc^{3m}/ x ∈ {a, b}* y la cantidad de b’s es par y m ≥ 0}
Solución:

S → aA/bbA
A → aA/bbA/C/ cccC
C→ cccC /ccc

5. Construye una gramática para el siguiente lenguaje:
L(G)={$b^n a^{n+1} c^{n+2}$ / n>=1}
Solución:

S → bBaAcc
B → a / bBaA
aAa → aaA
Ac → cc
Aa → aA

6. Construye una gramática para el siguiente lenguaje:
L(G)={$a^n c\, b^m$ / n>0 y m>=0}
Solución:

S → aAcB
A → aA /λ
B → bB/λ

7. Construye una gramática para el siguiente lenguaje:
L(G)={00X1 / X ∈ {0,1} }*
Solución:

S → 00X1
X → 0X/ 1X /λ

8. Construye una gramática para el siguiente lenguaje:
L(G)={X c^{3m} / X ∈ {a,b} y la cantidad de b’s es par y m>=0}*
Solución:

S → AC
A → aA /bbA/ λ
C → cccC/λ

9. Construye una gramática para el siguiente lenguaje:
L(G)={X/ X ∈ {0,1} y X contiene la subcadena 00 ó X contiene 11}*
Solución:

S → X00X/ X11X
X → 0X / 1X/ λ

*10. Construye una gramática que genere cadenas del alfabeto V=(a,b) que finalicen con **b** y que no tengan 2 **b**’s consecutivas.*
Solución:

S → b/ baAb/aAb
A → aA /abA/a/ λ

11. Construye una gramática que genere cadenas del alfabeto {a,b} que finalicen con ba
Solución:

S → Aba
A → aA/ bA /λ

12. Construye una gramática que genere cadenas del alfabeto V=(a,b) que no tengan 2 a’s consecutivas.
Solución:

S → abA/bAB
A → abA/bA/a/b/λ
A → baB

B → baB/b

13. Construye una gramática que genere cadenas del alfabeto V=(a,b) que genere el lenguaje:
$L(G)=\{a^n\ b^n$ /donde n sea múltiplo de 4}

Solución:

S → aaaaAbbbb
A → aaaaAbbbb/λ

14. Construye una gramática que genere cadenas del alfabeto V=(a,b) que genereun número par de aes

Solución:

S → aaAbB/ bBAaa/baAaB/abBaA
A → aaA/ λ
B → bB/ λ

15. Construye una gramática que genere cadenas del alfabeto V=(a,b) que genere un número par de aes y un número impar de bes.

Solución:

S → aaAbB
A → Aaa
B → bbB/ C
C →λ

16. Construye una gramática que genere cadenas del alfabeto V=(a,b) donde a sea siempre par

Solución:

S → AB/ BA
A → aaA/ bA/λ

17. Construye una gramática que genere cadenas del alfabeto V=(a,b) que genere un número impar de aes.

Solución:

S → AB/BA
A → aC
C → aaC/ λ
B → bB

18. Construye una gramática que genere cadenas del alfabeto V=(a,b) en la que cada instancia del símbolo a esté precedida y seguida de al menos una instancia de b. Ejm: bab, babab, bbabbbabbbb.

Solución:

S → BaB/ ABaB/ B
B → bB/ b
A → baA/ b

*19. Sea el vocabulario V=(a,b) y la expresión aa*bb*. Indicar el lenguaje que denota y algunas cadenas de dicho lenguaje.*

ab
aab
aaaab
abbbb
abb
aaab

Solución:

L={cadenas que comienzan por una a y continúan con varias o ninguna a, y siguen con una b y continúan con varias o ninguna b}

20. Construye una gramática tipo 0 y tipo 1 que genere el siguiente lenguaje:
$L(G)=\{c^{n+2}\ a^{n+1}\ c^{n}$ /n>=1}

Solución Tipo 0

S → ccAaBc

B → a/AaBc
aAa → Aaa
Ca → cc
aA → Aa

Solución Tipo 1

S → ccCaAD/ cccaac
C → cCAD
DA → XA
XA → XY
XY → AY
AY → AD
aA → aa
Da → DZ
DZ → DN
DN → NM
NM → ZD
ZD → aD
CA → ca
DD → cc
CD → cc

21. Construye una gramática que genere el siguiente lenguaje:
L(G)={$a^n b^m c$ /m>=0, n>=1, m múltiplo de 3, n par}

Solución:

S → ABc/ Ac
A → aaA/ aa
B → bbbB/ bbb

22. Construye una gramática que genere cadenas de V=(a,b) que no contengan 3 aes consecutivas

Solución:

S → AB/BA
A → aB/aaB/a
B → bB/bA/b

23. Construye una gramática que genere el siguiente lenguaje:
L(G)={$c^n a^{n+1} c^{n+2}$ /n>=1}

Solución:

S → cBaAcc
B → a/cBaA
aAa → aaA
Ac → cc
Aa → aA

24. Para cada una de los siguientes incisos, proporcione, si es posible, una gramática tipo 3 que defina el mismo lenguaje. Si en algún caso no es posible justifique.

a) ab^* / ba^*

Solución:

S → a / b / aA/ bB
A → bA/ b
B → aB/ a

b) $(a \mid b)^*c$

Solución:

S → c/ aS/ bS

c) b^*a / a^*b

Solución:

S→ a / b / bA/ aB
A→ bA/ a

B $\rightarrow$ aB/ b

25. Dada la siguiente gramática cuyas reglas de producción son de la forma:
P={S $\rightarrow$A, S $\rightarrow$ b B b, A $\rightarrow$ b C, A $\rightarrow$ **a** S B, B $\rightarrow$ **c** C, B $\rightarrow$ **a**, C$\rightarrow$ **c** C, C $\rightarrow$ b A}.
Se pide:

a) Razone si es posible que una cadena del lenguaje que define esta gramática empiece por el símbolo **a** y termine por el símbolo **b**. Argumente su respuesta.
b) Razone si es posible que una cadena del lenguaje que define esta gramática empiece por el símbolo **b** y termine por el símbolo **a**. Argumente su respuesta.

Solución:
a) Ninguna cadena del lenguaje definido por la gramática dada puede comenzar por el símbolo **a** y terminar por el símbolo **b**. Todas las reglas de producción de la gramática, excepto S $\rightarrow$ **b**B**b** y B $\rightarrow$ **a** terminan por A, B o C. Por lo tanto, comenzando la derivación utilizando S $\rightarrow$ A, las formas sentenciales que se obtienen siempre terminan por A, B o C, por lo que inevitablemente llegamos a cadenas que terminan por **a**. Teniendo en cuenta que comenzando la derivación utilizando S $\rightarrow$ **b**B**b** se obtienen cadenas que empiezan por **b**, se concluye que no es posible que una cadena del lenguaje dado cumpla con la condición antes mencionada.
b) Por otra parte, la cadena **bbababa** pertenece al lenguaje, luego es posible que una cadena del lenguaje que define esta gramática empiece por el símbolo **b** y termine por el símbolo **a**.
26. Dado el siguiente lenguaje, defina la gramática tipo 2 que lo genera.
L = {$a^m b^n c^k$ | m > n + k ; n, k >=0}
Solución:
Se puede desglosar de la siguiente manera:

L = {$a^m a^k a^n b^n c^k$ / m > 0; n, k >=0 }

S $\rightarrow$ AB
A$\rightarrow$ aA /a
B $\rightarrow$ aBc /C
C $\rightarrow$ aCb /ab

27. Dado el conjunto de palabras, determine la gramática y el lenguaje que las genera.
bcc, abccc, abbccccc, aabcccc, aabbcccccc, aaaaabccccccc,........
Se pide:

a) Definir un lenguaje mediante reglas de producción.
b) Demostrar con ejemplos las posibles derivaciones de S.

Solución:

V= {a, b, c}, L = {$a^n b^m c^{n+2m}$ | n >=0, m >= 1}

S $\rightarrow$ aSc | A
A $\rightarrow$ bAcc | bcc

28. Escribe la gramática tipo 1 que genere el siguiente lenguaje:

L = {$a^n b^m$ c / n>=1, m>= 0; n es múltiplo de 3 y m es par }

Solución:

S $\rightarrow$ ABc/ Ac
A $\rightarrow$ aaaA/aaa
B $\rightarrow$ bbB/bb

29. Escribe la gramática tipo 1 del alfabeto {a, b} que contengan como subcadena 3 aes consecutivas.
Solución:

S $\rightarrow$ aaaA/AaaaA/aaa/Aaaa
A $\rightarrow$ aA/a/ b/ bA

30. Dado el vocabulario V= {a, b}, escribe la gramática tipo 2 que genere el siguiente lenguaje:
L = { $a^n b^m$ | m <= n <= 2m ; n, m >=0}
Solución:

S $\rightarrow$ aSb | aaSb | ab

31. Escribe la gramática o reglas de producción tipo 0 que generen el siguiente lenguaje:

L = { $a^n b^m$ | n <> m ; n, m >0}

Donde el símbolo <> significa distinto.

Solución:

Podemos desglosar el lenguaje de la siguiente manera:

$$L = \{ a^n b^m \mid n > m ; n, m > 0\} \cup \{ a^n b^m \mid n < m ; n, m > 0\}$$

S → A / B
A → aAb / aA /aab
B → aBb /Bb /abb

Bibliografía

- Alfonseca, M., J. Sancho, M. A. Orga, *Teoría de Lenguajes, Gramáticas y Autómatas*, Promosoft, 1997.
- Cueva L., Juan Manuel, *Lenguajes Gramáticas y Autómatas,* Segunda Edición, 2001, Universidad de Oviedo.
- Díaz Víctor, Cañete José Miguel, *Lenguajes Formales y autómatas*, Universidad de Sevilla, 2006
- Ferreiro, Emilia y Margarita Gómez Palacio (Comp.) *Nuevas perspectivas sobre los procesos de lectura y escritura. Siglo XXI*. Buenos Aires.
- H. Contreras, *Los fundamentos de la gramática transformacional,* México, siglo XXI, 1971.
- Hopcroft, J.E., R. Motwani, J. D. Ullman, *Introducción a la teoría de autómatas, lenguajes y computación (2ª Edición)*, Prentice-Hall, 2000.
- J. Nivette: *Principios de gramática generativa.* Madrid, Fragua, 1973.
- Vidal Lamiquiz*, Lingüística Española.* Publicaciones de la Universidad de Sevilla, 1973.

Autores

Jhenny Castillo Tapia
Licenciada en Informática
Especialista en Docencia Universitaria
Docente en la carrera de Ingeniería Informática
Facultad del Gran Chaco
Yacuiba- Tarija- Bolivia

Ronald Yeber Cruz Delgado
Ingeniero en Informática
Consultor en Ingeniería Informática
Desarrollo de Software, Sistemas, Redes de Computadoras y Sistemas Telemáticos.
Tarija- Bolivia

Autores:
Lic. Jhenny Castillo Tapia
jhennyct@yahoo.es
Ing. Ronald Yeber Cruz Delgado
Bolivia 2008