Práctica 6:
Lenguajes Formales y Gramáticas

Cátedra: Lenguajes Formales y Computabilidad
Aulas Virtuales - Pandemia Covid19

Docentes en Práctica

- Alejandro Hernandez
- Natalia Colussi
- Valeria Perez Moguetta

Ejercicio 1:

1. Clasifique cada una de las siguientes gramáticas (dando su tipo más restrictivo):

- Debemos analizar cada producción de las gramáticas dadas y clasificarlas en
  - Tipo 3: REGULARES
  - Tipo 2: LIBRE DE CONTEXTO
  - Tipo 1: SENSIBLE al CONTEXTO
  - Tipo 0: ESTRUCTURADA POR FRASE

- L3 << L2 << L1 << L0

Las gramáticas regulares son más estrictas que las libre de contexto, y éstas últimas más estrictas que las sensibles al contexto, y las sensible al contexto más estrictas que las estructuradas por frase.

Ejercicio 1:

Jerarquía de Chomsky

- Lo que buscamos en este ejercicio es detectar mediante las reglas de producción de cada gramática a qué tipo pertenece.

REGULAR	$A \to a,\ A \to aB$ ó $A \to \lambda$ donde $A, B \in N$ y $a \in T$;
INDEPENDIENTE DEL CONTEXTO	$A \to \delta$ donde $A \in N$ y $\delta \in (N \cup T)^*$
SENSIBLE AL CONTEXTO	$\alpha A \beta \to \alpha \delta \beta$ donde $A \in N,\ \alpha, \beta \in (N \cup T)^*$ y $\delta \in (N \cup T)^+$; $\sigma \to \lambda$ siendo σ el estado inicial
ESTRUCTURADA POR FRASES	$\alpha \to \delta$ donde $\alpha \in (N \cup T)^* - T^*$ y $\delta \in (N \cup T)^*$

Ejercicio 1:

(a) $T = \{a, b\}$, $N = \{\sigma, A\}$, símbolo inicial σ, y producciones:

i. $\sigma \to b\sigma$

ii. $\sigma \to aA$

iii. $\sigma \to b$

iv. $A \to a\sigma$

v. $A \to bA$

vi. $A \to a$

Ninguna de las reglas viola la estructura de las producciones de las gramáticas regulares, por lo tanto, la gramática dada en el ítem (a) es REGULAR

REGULAR	$A \to a,\ A \to aB$ ó $A \to \lambda$ donde $A, B \in N$ y $a \in T$;

Ejercicio 1:

(d) $T = \{a, b, c\}$, $N = \{\sigma, A, B\}$, símbolo inicial σ, y producciones:

i. $\sigma \to BAB$

ii. $\sigma \to ABA$

iii. $A \to AB$

iv. $B \to BA$

v. $A \to aA$

vi. $A \to ab$

vii. $B \to b$

La gramática es libre de contexto todas las producciones se ajustan a las restricciones de esta categoría, además las reglas (i) a (iv) no son regulares.

REGULAR	$A \to a,\ A \to aB$ ó $A \to \lambda$ donde $A, B \in N$ y $a \in T$;
INDEPENDIENTE DEL CONTEXTO	$A \to \delta$ donde $A \in N$ y $\delta \in (N \cup T)^*$

Ejercicio 2:

2. Dé una derivación de las siguientes cadenas en las gramáticas especificadas:

(a) $bbabbab$, gramática 1a

(b) $abab$, gramática 1b

(c) $aabaab$, gramática 1c

(d) $abbabb$, gramática 1d

(b) $T = \{a, b, c\}$, $N = \{\sigma, A, B\}$, símbolo inicial σ, y producciones:

i. $\sigma \to AB$

ii. $AB \to BA$

iii. $A \to aA$

iv. $B \to Bb$

v. $A \to a$

vi. $B \to b$

2b. σ -> (regla i) AB -> (regla iii) aAB -> (regla iv) aABb -> (regla ii) aBAb -> (regla v) aBab -> (regla vi) abab

Ejercicio 3:

3. Muestre que la cadena *abbbabaaba* no está en el lenguaje generado por la gramática $G = (N, T, P, \sigma)$, donde $N = \{\sigma, A, B\}$, $T = \{a, b\}$, símbolo inicial σ, y producciones:

$$\sigma \to aaBA,\ \sigma \to ABB,\ A \to aaB,\ A \to \lambda,\ aBa \to A,\ Aaa \to B,\ B \to AabaB,\ B \to bbb$$

Sugerencia: Piense en toda derivación como una secuencia de cadenas sobre el alfabeto $N \cup T$ y halle un invariante, verificado por la cadena σ, que se preserve en la aplicación de todas las reglas de producción pero no sea verificado por la cadena propuesta.

- Esto nos debe “sonar” familiar de haberlo hecho en algún parcial o en alguna práctica pasada. Se acuerdan en cual?
- La idea es pensar y analizar de qué forma podemos demostrar que la cadena en cuestión no pertenece al conjunto de palabras posibles sobre los conjuntos N U T.

Ejercicio 3.

La cadena $\overbrace{abbbabaaba}^{\boxed{w}}$ no está en el lenguaje generado por la gramática dada.

Para realizar esta prueba utilizaremos la TEORÍA DE CONJUNTOS INDUCTIVOS.

$$L(G) = \{\alpha \in T^* \mid \sigma \overset{*}{\Rightarrow} \alpha\} \subseteq (N \cup T)^*$$

Las reglas de producción son las reglas de construcción de los elementos del conjunto.

Para probar que no pertenece al conjunto $L(G)$ tenemos que encontrar una propiedad que cumplan los elementos de $L(G)$ pero que no cumple la cadena $\boxed{w}$.

Buscamos una propiedad...

$\sigma \to aaBA$ | $A \to aaB$ | $B \to bbb$
$\sigma \to ABB$ | $A \to \lambda$ | $B \to Aaba$
| $Aaa \to B$ | $aBa \to A$

Agregamos o eliminamos letras "a" en una cantidad par.

$w = abbbabaaba$ tiene una cantidad impar de "a"

Propiedad

$P(z)$: z tiene una cantidad par de letras a

$$\left(\forall z : L(G) \bullet P(z)\right)$$

Para probar esta propiedad aplicaremos inducción en la cantidad de pasos en la derivación de una cadena

CASO BASE : n = 1

Analizamos las reglas de producción.

1 sólo paso:

aaBA	#a es par (2)
ABB	#a es par (0)
aaB	#a es par (2)
λ	#a es par (0)
B	#a es par (0)
bbb	#a es par (0)
Aaba	#a es par (2)
A	#a es par (0)

∴ Se cumple para el caso base.

CASO INDUCTIVO : $n \Rightarrow n+1$

Supongamos

$P(z)$ se cumple con n pasos de derivación

Esto es que, la cantidad de letras "a" para z es par, $\#_a z = 2k$ con $k \in \mathbb{N}$

CASO INUCTIVO : $n \Rightarrow n+1$

Demostrar que:

$P(z')$ se cumple con $n+1$ pasos de derivación donde z' se obtuvo de aplicar alguna de las reglas sobre z.

CASO INUCTIVO : $n \Rightarrow n+1$

Analicemos:

1º) $\boxed{A \to aaB}$, aplicando resulta z'

$$\#_a z' \underset{\text{prop. conteo}}{=} \#_a z + 2 \underset{HI}{=} 2k+2 = 2 \times (k+1)$$

donde también resulta un nro par.

2º) $\boxed{A \to \lambda}$ resultando z'

$$\#_a z' = \#_a z \underset{HI}{=} 2k.$$ que resulta ser un nro par.

... COMPLETAR.

$\therefore$ Propiedad P se cumple para todo $z \in L(G)$.

Sea cadena w no satisface P por lo tanto podemos afirmar que

$$w \notin L(G).$$

Ejercicio 4:

4. Dé una gramática del tipo pedido que genere los siguientes lenguajes:

- Gramática regular:

  (a) Cadenas sobre el alfabeto $\{a, b\}$ que comiencen con a.

  (b) Cadenas sobre el alfabeto $\{a, b\}$ que contengan exactamente una a y terminen con al menos una b.

  (c) Cadenas sobre el alfabeto $\{a, b\}$ que terminen con ba.

  (d) Cadenas sobre el alfabeto $\{a, b\}$ que contengan ba.

  (e) Cadenas sobre el alfabeto $\{a, b\}$ que no terminen con ab.

4a. G = ({C}, {a, b}, P, σ), donde P = {σ -> aC, C -> aC, C -> bC, C->λ}

- Gramática independiente de contexto:
  (f) Cadenas sobre el alfabeto $\{a, b\}$ de la forma $a^n b^n$ para $n \geq 0$.
  (g) Cadenas sobre el alfabeto $\{true, p, q, \wedge, \vee, \neg, (,)\}$ definido inductivamente como el menor conjunto tal que:
    i. $true \in \mathcal{L}$
    ii. $p \in \mathcal{L}$
    iii. $q \in \mathcal{L}$
    iv. Si $A \in \mathcal{L}$ entonces $(\neg A) \in \mathcal{L}$.
    v. Si $A, B \in \mathcal{L}$ entonces $(A \wedge B) \in \mathcal{L}$.
    vi. Si $A, B \in \mathcal{L}$ entonces $(A \vee B) \in \mathcal{L}$.
  (h) Cadenas palíndromes[1] sobre el alfabeto $\{a, b\}$.
  (i) $\{a^n b^n c^k \mid n, k \in \mathbb{N}\}$
  (j) Cadenas sobre el alfabeto $\{a, b\}$ que empiezan y terminan con el mismo símbolo.
  (k) Cadenas sobre el alfabeto $\{a, b\}$ de longitud impar.
  (l) Cadenas sobre el alfabeto $\{a, b, c\}$ de longitud impar y cuyo símbolo central es c.
  (m) $\{a^n b^{2n} c^m \mid n, m \in \mathbb{N}_0\}$
  (n) $\{a^n b^m c^m \mid n, m \in \mathbb{N}_0\}$

Ejercicio 5:

5. Sea $\mathcal{L}$ el conjunto de cadenas sobre $\{a, b\}$ que contienen la misma cantidad de símbolos a y b. Analice si cada una de las siguientes gramáticas genera $\mathcal{L}$. En caso negativo, dé un contraejemplo (es decir, una cadena que sea generada por la gramática pero no esté en $\mathcal{L}$, o una cadena que esté en $\mathcal{L}$ pero no sea generada por la gramática). En todas las gramáticas el símbolo inicial es S.

   (a) $S \to aSb \mid bSa \mid \lambda$

   (b) $S \to aSb \mid bSa \mid \lambda \mid SS$

   (c) $S \to aB \mid bA \mid \lambda$, $B \to b \mid bA$, $A \to a \mid aB$

   (d) $S \to aSb \mid baS \mid abS \mid bSa \mid \lambda$

   (e) $S \to aB \mid bA$, $A \to a \mid Sa$, $B \to b \mid SB$

Ejercicio 5:

$\mathcal{L}_e = \{ w \in \{a, b\}^* \cdot \#_a w = \#_b w \}.$

Tenemos que determinar si las gramaticas dadas generan $\mathcal{L}_e$.

$$\left(L(G) = \mathcal{L}_e \right) \equiv L(G) \subseteq \mathcal{L}_e \wedge \mathcal{L}_e \subseteq L(G)$$

(a) $S \to aSb$ FALSO!

$S \to bSa$ ✓

$S \to \lambda$.

La cadena abba $\in \mathcal{L}_e$ pero no es generada por G.

abbaba tampoco es generado por G.

Ejercicio 6:

Definición 2 Una gramática se dice que está en *forma normal de Chomsky* sii toda producción es de la forma:

$$A \to BC, \ A \to a \text{ ó } S \to \lambda$$

donde $A, B, C \in N$, $a \in T$, S es el no terminal inicial y B y C son distintos de S.

6. Convierta a una gramática equivalente que esté en forma normal de Chomsky:

- $S \to xSy$
- $S \to wNz$
- $N \to S$
- $N \to \lambda$

Cada regla se puede tratar independientemente, por ejemplo:

S -> wNz se podría transformar en S -> WM, M -> NZ, W -> w, N -> n, Z -> z

Ejercicio 6:

6. Convierta a una gramática equivalente que esté en forma normal de Chomsky:

- $S \to xSy$
- $S \to wNz$
- $N \to S$
- $N \to \lambda$

La dificultad radica en que, en algunas reglas, S ocurre del lado derecho. Esto lo tenemos que eliminar “previamente”, pero para esto <u>**sí necesitamos**</u> mirar reglas en conjunto con otras reglas, no puede ser independiente.

Por ejemplo, las reglas N -> S y S -> wNz se pueden reemplazar por N -> wNz (y luego se puede aplicar la técnica de la slide anterior).

Otro caso: las reglas S -> xSy y S -> wNz se podrían reemplazar por S -> xwNzy

Ejercicio 7:

(a) Se puede probar que el lenguaje $\mathcal{L} = \{a^n b^n c^n \mid n \in \mathbb{N}\}$ no es independiente de contexto. Usando este resultado, probar que el conjunto de los lenguajes independientes de contexto no es cerrado bajo la operación de intersección, hallando dos lenguajes que sean independientes de contexto cuya intersección sea $\mathcal{L}$.

INDEPENDIENTE DEL CONTEXTO	$A \to \delta$ donde $A \in N$ y $\delta \in (N \cup T)^*$

a) $\mathcal{L} = \{a^n b^n c^n \circ n \in \mathbb{N}\}$

$\mathcal{L}$ NO es independiente del contexto

(Puntualmente $\mathcal{L}$ es SENSIBLE AL CONTEXTO)

Ejercicio 7:

Demostrar

L_{e1} es Independiente del contexto

L_{e2} es " " "

$$\boxed{L_{e1} \cap L_{e2} \text{ \underline{no} es Indep. del Contexto}}$$

Es decir que <u>no</u> es cerrado para la intersección

Ejercicio 7:

Para poder encontrar a $\mathcal{L}_{e1}$ y $\mathcal{L}_{e2}$ veamos cómo es $\mathcal{L}_e$.

$$\mathcal{L}_e = \{abc, aabbcc, aaabbbccc, aaaabbbbcccc, \ldots\}$$

Pensamos en dos lenguajes que contengan a $\mathcal{L}_e$.

$\mathcal{L}_{e1} = \{a^n b^n c^k \bullet n, k \in \mathbb{N}\}$ es LIBRE CONTEXTO

$\mathcal{L}_{e2} = \{a^k b^n c^n \bullet n, k \in \mathbb{N}\}$ es LIBRE DE CONTEXTO

Ejercicio 7:

Si $L_c = L_{c1} \cap L_{c2}$ la intersección de lenguajes libres de contexto fuese cerrada entonces L_c sería libre de contexto. Pero L_c es sensible de contexto.

Luego la $\cap$ no es cerrada para los L_c.

LIBRE CONTEXTO

Ejercicio 8:

8. Se puede probar que la intersección de un lenguaje independiente del contexto y uno regular es independiente de contexto. Usar este resultado y el enunciado del ejercicio 7a para probar que el lenguaje $\{w \in \{a, b, c\}^* \mid N_a(w) = N_b(w) = N_c(w)\}$ no es independiente del contexto.

Notación: $N_x(w)$ denota la cantidad de ocurrencias del símbolo x en la cadena w.

Sabemos que:

1) C = $\{a^n b^n c^n\}$ no es independiente de contexto
2) Si A es IC y B es Reg. => A intersección B es IC

Sea A' el lenguaje del enunciado. Si encontramos un B regular, tal que la intersección de A' con B es C, queda demostrado que A' no es IC (por contrarrecíproca de 2, o absurdo de suponer A' IC).

Hint: pensemos en que los elementos de A' solo cumplen con una propiedad (cantidad de ocurrencias) versus los elementos de C que cumplen con dos (ocurrencias y además ordenadas).

Ejercicio 9:

9. Probar que si $\mathcal{L}$ es un lenguaje independiente de contexto, entonces $\mathcal{L}^R$ también es independiente de contexto.

$\mathcal{L}$ es independiente del contexto

$\mathcal{L}^R$ también es independiente del contexto

Dem)

$\mathcal{L}^R$ denota al lenguaje reverso de $\mathcal{L}$

$\mathcal{L}^R = \{ w^R \in \Sigma^* \cdot w \in \mathcal{L} \}$

donde

$w = a_1 a_2 a_3 \dots, a_n \in \mathcal{L}$

$w^R = a_n a_{n-1}, \dots, a_2, a_1$

¿Qué propiedades cumple los elementos inversos?

Ejercicio 10:

10. Sea la gramática con producciones:

$$S \to aS \mid Sb \mid a \mid b$$

y símbolo inicial S. Pruebe, por inducción sobre el número de pasos en la derivación, que toda cadena generada por la gramática no contiene la subcadena ba.

Caso base: derivación de 1 paso.

Subcaso usando regla S -> a. Trivial

Subcaso usando regla S -> b. Trivial

Ejercicio 10:

10. Sea la gramática con producciones:

$$S \to aS \mid Sb \mid a \mid b$$

y símbolo inicial S. Pruebe, por inducción sobre el número de pasos en la derivación, que toda cadena generada por la gramática no contiene la subcadena ba.

Caso inductivo: supongo que toda derivación de n pasos genera cadenas que no contienen la subcadena ba. Probaremos que toda derivación de n+1 pasos genera cadenas que no contienen la subcadena ba.

Subcaso “paso 1 es S -> aS”. Después de ese primer paso, hay n pasos a partir de la S obtenida, que generan la cadena w. Por ende con los n+1 pasos se genera aw. Por HI w no contiene la subcadena ba. Tampoco se puede formar la subcadena ba usando la a y el primer carácter de w. Ergo ba no ocurre en aw.

Ejercicio 11:

11. Considere la gramática libre de contexto $G = (\{E, V\}, \{+, *, (,), x, y, z\}, P, E)$ donde P consiste de las producciones:

$$E \to E + E \mid E * E \mid V \mid (E)$$

$$V \to x \mid y \mid z$$

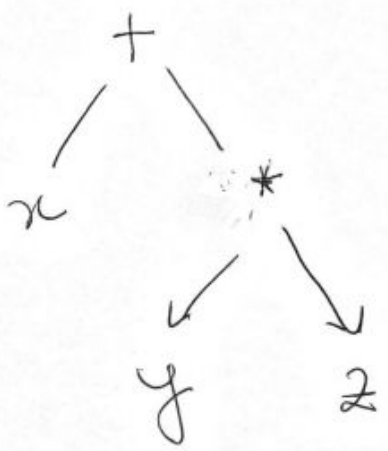
   (a) Pruebe que la gramática es ambigua.

   (b) Dé una gramática libre de contexto equivalente que resuelva la ambigüedad mediante la precedencia usual entre la multiplicación y la suma.

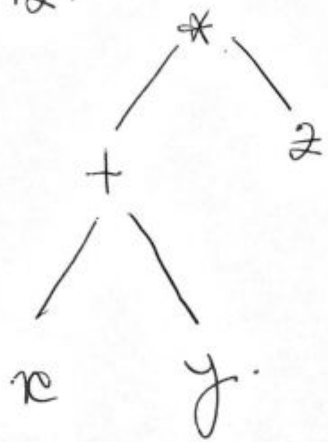
Ejercicio 11:

$x + y * z$ ¿Cuál es el árbol de parseo?

T_1:



T_2:



La gramática es ambigua, ya que tiene mas de un árbol de parseo.