Teoría de Autómatas y Lenguajes Formales

Boletín de Autoevaluación 6: **¿Cómo obtener Formas Normales de una GCL?.**

1. Objetivos.

El objetivo de este boletín es ilustrar cómo proceder para obtener gramáticas de contexto libre en formas normales, en concreto la Forma Normal de Chosmky la Forma Normal de Greibach, y, además, proporcionar la solución a alguno de los problemas propuestos en el boletín para que podáis comprobar si habéis aplicado bien este método.

2. Idea Principal.

La jerarquía de Chomsky indica cuál debe ser el formato general que cumplan las producciones de una gramática para generar un lenguaje de una determinada clase. Pero ese formato es tan general que da lugar a múltiples formas gramaticales posibles para las producciones. Especialmente, en el caso de las GCLs.

Una forma normal es un intento de estandarizar las producciones y conseguir que todas tengan una apariencia similar. Esa sería la primera idea. Pero tienen más utilidades.

En la asignatura se estudian dos formas normales distintas para las GCLs: la forma normal de Chomsky (FNCh) y la forma normal de Greibach (FNG).

En ambos casos se debe partir de un GCL *ya simplificada*.

Ilustraremos la explicación con la siguiente gramática:

$$\begin{array}{l} S \rightarrow A \mid AA \mid AAA \\ A \rightarrow ABa \mid ACa \mid a \\ B \rightarrow ABa \mid Ab \mid \lambda \\ C \rightarrow Cab \mid CC \\ D \rightarrow CD \mid Cd \mid CEa \\ E \rightarrow b \end{array}$$

que tal y como se vio en el boletín de autoevaluación 5, es equivalente a la gramática simplificada

$$\begin{array}{l} S \rightarrow ABa \mid Aa \mid a \mid AA \mid AAA \\ A \rightarrow ABa \mid Aa \mid a \\ B \rightarrow ABa \mid Aa \mid Ab \end{array}$$

Servirá como ejemplo para ilustrar primero la Forma Normal de Chomsky y, después, la Forma Normal de Greibach.

3. Forma Normal de Chomsky.

La teoría define así la FNCh:

"Todo LCL L, $\lambda \notin L$, se puede generar por una GCL en la que todas sus producciones tienen el siguiente formato:

$$A \rightarrow BC$$
$$A \rightarrow a$$

donde A, B, C son símbolos auxiliares y a es un símbolo terminal".

Para entender el porqué de esta forma normal, además de por sus aplicaciones prácticas (por ejemplo, para ver si el lenguaje generado por una GCL es o no es finito), hay que recordar que Chomsky es lingüista y. aún más, es el padre de la teoría *generativa*, que dice que el ser humano nace con unas estructuras gramaticales innatas. Llegó a esta conclusión tras analizar la estructura similar de muchos lenguajes que, aparentemente, son muy diferentes. Pero casi todos comparten frases con una estructura de $\langle SUJETO\ PREDICADO\rangle$ (hmmm... un par de auxiliares) y, a su vez, cada una de estas partes comparte un estructura de tipo $\langle NÚCLEO\ COMPLEMENTOS\rangle$ (¡vaya! otro par de auxiliares..) y si se sigue el análisis se ve que, por ejemplo, el $NÚCLEO$ del $PREDICADO$ es un $verbo$ (caramba, vamos que $NÚCLEO \rightarrow verbo$, o sea que un auxiliar se deriva en un terminal... curioso).

Bien, sea como sea, para obtener la FNCh hay que seguir dos pasos. Primero, si se observan los dos formatos permitidos, se ve que no se pueden mezclar símbolos auxiliares y símbolos terminales en una misma producción. Para conseguir eso, se definen tantos auxiliares como haga falta para substituir a los terminales que aparezcan mezclados con auxiliares en los consecuentes de las producciones. En el ejemplo, el efecto será:

Gramática simplificada

$$\begin{array}{l} S \rightarrow ABa \mid Aa \mid a \mid AA \mid AAA \\ A \rightarrow ABa \mid Aa \mid a \\ B \rightarrow ABa \mid Aa \mid Ab \end{array}$$

Gramática "pseudo-Chomsky"

$$\begin{array}{l} S \rightarrow ABC_a \mid AC_a \mid a \mid AA \mid AAA \\ A \rightarrow ABC_a \mid AC_a \mid a \\ B \rightarrow ABC_a \mid AC_a \mid AC_b \\ C_a \rightarrow a \\ C_b \rightarrow b \end{array}$$

Como se ve, puesto que a y b aparecían mezclados con auxiliares en distintas producciones, por ejemplo $(S \rightarrow ABa)$ ó $(B \rightarrow Ab)$, se han definido dos nuevos auxiliares, C_a y C_b cuyas únicas producciones son $(C_a \rightarrow a)$ y $(C_b \rightarrow b)$. Al substituir, por ejemplo $(S \rightarrow ABC_a)$ ó $(B \rightarrow AC_b)$, desaparecen las "mezclas".

Ojo, que en producciones como $(S \rightarrow a)$ ó $(A \rightarrow a)$ *no hay que substituir*, ya están en forma normal de Chomsky.

Una vez obtenida esta forma "pseudo-Chomsky", se aplica el segundo paso, que tiene que ver con que el único formato permitido en producciones de auxiliares es *por parejas*. Por lo tanto, en todas las producciones que aparezcan más de dos símbolos auxiliares hay que conseguir una agrupación que permita cumplir con esta regla. En el ejemplo se podría pensar en lo siguiente:

Gramática "pseudo-Chomsky"

$$\begin{array}{l} S \rightarrow A\mathbf{BC_a} \mid AC_a \mid a \mid AA \mid A\mathbf{AA} \\ A \rightarrow A\mathbf{BC_a} \mid AC_a \mid a \\ B \rightarrow A\mathbf{BC_a} \mid AC_a \mid AC_b \\ C_a \rightarrow a \\ C_b \rightarrow b \end{array}$$

Gramática en FNCh

$$\begin{array}{l} S \rightarrow AX \mid AC_a \mid a \mid AA \mid AY \\ A \rightarrow AX \mid AC_a \mid a \\ B \rightarrow AX \mid AC_a \mid AC_b \\ C_a \rightarrow a \\ C_b \rightarrow b \\ X \rightarrow BC_a \\ Y \rightarrow AA \end{array}$$

Se ha agrupado el par de auxiliares BC_a y se introduce un nuevo auxiliar cuya única producción sea $(X \rightarrow BC_a)$. Con substituir todas las apariciones del par BC_a por el nuevo auxiliar X se eliminan casi todos los trios de auxiliares. Para eliminar el trio AAA se hace algo similar, introduciendo el auxiliar Y. Nótese que sólo se utiliza Y para transformar $(S \rightarrow AAA)$ en $(S \rightarrow AY)$; sobre la producción $(S \rightarrow AA)$ *no se aplica la substitución* porque ya está en forma de Chomsky.

4. Forma Normal de Greibach.

La teoría dice:

> *"Todo LCL L, $\lambda \notin L$, se puede generar por una GCL en la que todas sus producciones tienen el siguiente formato:*
>
> $$A \rightarrow a\alpha$$
>
> *donde A es un símbolo auxiliar, a es un símbolo terminal y α es una cadena de Σ_A^*".*

es decir, la forma normal de Greibach consiste en hacer que el consecuente de todas las producciones comience con un terminal y que éste vaya seguido por una subcadena (que puede ser λ) de auxiliares. La justificación de esta forma normal queda patente al estudiar la relación entre esta y los autómatas de pila, como veremos más adelante. Por ahora nos conformaremos con el proceso para obtenerla :-)

Bueno, en este momento dejo el estilo "ceremonioso" para asumir (a) que has intentado hacer alguna vez una FNG y (b) en cuanto te han contado lo de *el orden de los símbolos auxiliares* has pegado un bufido y has gritado (mentalmente (o no)) “¡¡¿Por qué?!!”. Vamos, que te suena el tema y por eso has acudido al boletín de autoevaluación.

Sí, es cierto. Una de las primeras cosas que decimos los profesores de TALF es eso de: “...nótese que el primer paso para la obtención de la Forma de Greibach de una gramática que ya está en la forma normal de Chomsky, consistirá en el establecimiento de un orden entre los símbolos auxiliares de la gramática a fin de ir substituyéndolos en las producciones en ese mismo orden para obtener la forma normal”.

Nada, con un poco de música podría ser una sinfonía. Vamos a ir digiriendo el párrafo.

Partimos de una gramática en forma de Chomsky. O, ya puestos, yo recomiendo[1] la forma “pseudo-Chomsky”: es igual de válida, pero seguro que tiene menos (o igual) número de auxiliares que la FNCh. Y, como veremos, cuantos más auxiliares entren en juego, más puede complicarse el proceso.

Fijaos que si partimos de una “pseudo-Chomsky” o una FNCh, seguro que ya disponemos de producciones que están en forma de Greibach, porque producciones como

$$A \rightarrow a$$

son comunes a ambas formas normales. Y entra en juego el “orden”. ¿Para qué establecerlo? El caso es que si lo hacemos bien la forma normal se calcula automáticamente. Retomemos el ejemplo: se había obtenido la forma “pseudo-Chomsky” de la gramática original como

$$\begin{array}{l} S \rightarrow ABC_a \mid AC_a \mid a \mid AA \mid AAA \\ A \rightarrow ABC_a \mid AC_a \mid a \\ B \rightarrow ABC_a \mid AC_a \mid AC_b \\ C_a \rightarrow a \\ C_b \rightarrow b \end{array}$$

En esa gramática ya hay auxiliares con sus producciones en forma de Greibach, C_a y C_b. Y otros que no, evidentemente. Nos fijamos, por ejemplo, en las producciones de B,

$$B \rightarrow ABC_a \mid AC_a \mid AC_b$$

y podríamos pensar los siguiente: “Vaya, si el símbolo A ya estuviera en FNG, al substituir A por sus producciones B quedaría *automáticamente* en forma de Greibach”. Y fíjaos, además, que en teoría se ha visto un lema relacionado con el tema y que ahora se puede “amortizar”: el *lema de substitución*,

Sea $G = \langle \Sigma_A, \Sigma_T, P, S \rangle$ *una GCL. Sea* $(A \rightarrow \alpha_1 B \alpha_2) \in P$ *y sea* $(B \rightarrow \beta_1 \mid \beta_2 \mid \ldots \mid \beta_r)$ *el conjunto de todas las producciones de G que tienen al símbolo auxiliar B como antecedente. Sea* $G' = \langle \Sigma_A, \Sigma_T, P', S \rangle$ *la GCL tal que P′ se ha obtenido del conjunto P al eliminar de este la producción* $(A \rightarrow \alpha_1 B \alpha_2)$ *y añadir las producciones* $(A \rightarrow \alpha_1 \beta_1 \alpha_2 \mid \alpha_1 \beta_2 \alpha_2 \mid \ldots \mid \alpha_1 \beta_r \alpha_2)$*, entonces*

$$L(G) = L(G').$$

[1] Salvo que en un enunciado se os pida la FNCh explícitamente, claro está.

que nos indica que si un símbolo auxiliar es substituido por sus producciones, allá donde aparezca, la gramática obtenida es equivalente.

Puedo aprovecharme de este lema, entonces, para intentar la siguiente estrategia: en todas las producciones que no estén en forma de Greibach me fijo en el primer símbolo de los consecuentes:

$$
\begin{array}{l}
S \to \mathbf{A}BC_a \mid \mathbf{A}C_a \mid a \mid \mathbf{A}A \mid \mathbf{A}AA \\
A \to \mathbf{A}BC_a \mid \mathbf{A}C_a \mid a \\
B \to \mathbf{A}BC_a \mid \mathbf{A}C_a \mid \mathbf{A}C_b \\
C_a \to a \\
C_b \to b
\end{array}
$$

porque eso tiene que ser útil para poder establecer el *orden de substitución* que, al fin y al cabo, lo tengo que establecer para ir substituyendo símbolos y que las producciones cumplan la FNG. Si todas las producciones de B, por ejemplo, que no están en forma de Greibach tienen el consecuente comenzando por A, *me conviene establecer la Forma de Greaibach de A antes que la de B*, lo que se suele expresar como $A > B$. Si todas las producciones de S, por ejemplo, que no están en forma de Greibach tienen el consecuente comenzando por A, *me conviene establecer la Forma de Greaibach de A antes que la de S*, $A > S$.

A estas alturas, si has pillado el razonamiento tendrías que estar levantando ya la mano y decir: "Sí, vale pero ¿y qué hacemos con las producciones de A? ¿Me vas a decir que substituimos A antes que A o que $A > A$?". Si te ha surgido la pregunta, pasa al párrafo siguiente :-). Si no, vuelve al anterior ;-)

¡¡Bienvenido al apasionante mundo de la *recursividad a izquierdas*!! Efectivamente, en las producciones de A que no están en forma de Greibach, los consecuentes comienzan con el símbolo A. ¿Cómo se soluciona?

> *Sea $G = \langle \Sigma_A, \Sigma_T, P, S \rangle$ una GCL y sea $A \to A\alpha_1 \mid A\alpha_2 \mid \ \ldots \ \mid A\alpha_r$ el conjunto de todas las producciones de P que tienen al símbolo A como antecedente y que son recursivas a izquierdas. Sea $A \to \beta_1 \mid \beta_2 \mid \ \ldots \ \mid \beta_m$ el resto de producciones de P que tienen al símbolo A como antecedente.*
>
> *Sea $G' = \langle \Sigma'_A, \Sigma_T, P', S \rangle$ la GCL obtenida de G tal que $\Sigma'_A = \Sigma_A \cup \{B\},\ B \notin \Sigma_A$, y tal que en P′ se han eliminado todas las producciones que tienen al símbolo A como antecedente y en su lugar se han añadido las siguientes producciones:*
>
> 1. $A \to \beta_i \mid \beta_i B,\ 1 \le i \le m$,
> 2. $B \to \alpha_i \mid \alpha_i B,\ 1 \le i \le r$.
>
> *Entonces L(G) = L(G′)*

Este es el lema de la *eliminación de la recursividad a izquierdas*. Para hacer lo que pide el lema, dividimos las producciones de A (seguimos con el mismo ejemplo), en dos tipos

$$A \to \mathbf{ABC_a} \mid \mathbf{AC_a} \mid a$$

Las que están marcadas en negrita, son las que presentan recursividad. Para aplicar el lema, dejamos en A sólo las producciones *sin recursividad* tal cual están y acabadas en un nuevo símbolo auxiliar, X,

$$A \to a \mid aX$$

Y, por supuesto, si introducimos un nuevo auxiliar, hay que definir sus producciones. Volvemos a fijarnos en las producciones originales de A recursivas a izquierdas... la verdad es que no estaban tan mal... lástima esa A que provoca la recursividad... pues ¡la quitamos! Y definimos X con lo que quede de los consecuentes, tal cual y acabados en X:

$$X \to BC_a \mid C_a \mid BC_aX \mid C_aX$$

Aparece una producción unitaria (que elimino para volver a obtener una gramática simplificada, substituyo C_a por a). La nueva gramática es:

$$\begin{array}{l} S \to ABC_a \mid AC_a \mid a \mid AA \mid AAA \\ A \to a \mid aX \\ X \to BC_a \mid a \mid BC_aX \mid C_aX \\ B \to ABC_a \mid AC_a \mid AC_b \\ C_a \to a \\ C_b \to b \end{array}$$

Y volvemos a intentar establecer el orden. Nos fijamos en el primer símbolo de cada consecuente de una producción que no esté en FNG,

$$\begin{array}{l} S \to \mathbf{A}BC_a \mid \mathbf{A}C_a \mid a \mid \mathbf{A}A \mid \mathbf{A}AA \ (\Rightarrow A > S) \\ A \to a \mid aX \\ X \to \mathbf{B}C_a \mid a \mid \mathbf{B}C_aX \mid \mathbf{C_a}X \ (\Rightarrow B > X, C_a > X) \\ B \to \mathbf{A}BC_a \mid \mathbf{A}C_a \mid \mathbf{A}C_b \ (\Rightarrow A > B) \\ C_a \to a \\ C_b \to b \end{array}$$

Se han obtenido “pistas” para determinar el orden: $A > S \wedge B > X \wedge C_a > X \wedge A > B$. Cualquier ordenación de los símbolos que respete estos requerimientos será válida y permite calcular la FNG. Por ejemplo, con el orden

$$C_b > C_a > A > S > B > X$$

se irá obteniendo:

1. Producciones de C_b,

$$C_b \to b$$

2. Producciones de C_a,

$$\begin{array}{l} C_b \to b \\ C_a \to a \end{array}$$

3. Producciones de A,

$$\begin{array}{l} C_b \to b \\ C_a \to a \\ A \to a \mid aX \end{array}$$

4. Producciones de S, substituimos A por los consecuentes de sus producciones,

$$\begin{array}{l} C_b \to b \\ C_a \to a \\ A \to a \mid aX \\ S \to aBC_a \mid aXBC_a \mid aC_a \mid aXC_a \mid a \mid aA \mid aXA \mid aAA \mid aXAA \end{array}$$

5. Producciones de B, substituimos A por los consecuentes de sus producciones,

$$
\begin{array}{l}
C_b \to b \\
C_a \to a \\
A \to a \mid aX \\
S \to aBC_a \mid aXBC_a \mid aC_a \mid aXC_a \mid a \mid aA \mid aXA \mid aAA \mid aXAA \\
B \to aBC_a \mid aXBC_a \mid aC_a \mid aXC_a \mid aC_b \mid aXC_b
\end{array}
$$

6. Producciones de X, substituimos B y C_a por los consecuentes de sus producciones,

$$
\begin{array}{l}
C_b \to b \\
C_a \to a \\
A \to a \mid aX \\
S \to aBC_a \mid aXBC_a \mid aC_a \mid aXC_a \mid a \mid aA \mid aXA \mid aAA \mid aXAA \\
B \to aBC_a \mid aXBC_a \mid aC_a \mid aXC_a \mid aC_b \mid aXC_b \\
X \to aBC_aC_a \mid aXBC_aC_a \mid aC_aC_a \mid aXC_aC_a \mid aC_bC_a \mid aXC_bC_a \mid a \mid \\
\qquad aBC_aC_aX \mid aXBC_aC_aX \mid aC_aC_aX \mid aXC_aC_aX \mid aC_bC_aX \mid aXC_bC_aX \mid aX
\end{array}
$$

La gramática está ya en Forma Normal de Greibach.

4.1. Cuestiones relacionadas con la obtención de la FNG.

4.1.1. ¿Qué pasa si al simplificar la gramática veo que *S* es anulable?

O, lo que es lo mismo, que $\lambda \in L(G)$. De eso tendríamos que darnos cuenta al eliminar las producciones vacías. Ya se comentó en el boletín de autoevaluación número 5: en principio, nos "apuntamos" que $S \to \lambda$ debería ser una producción de nuestra gramática, pero seguimos trabajando con la gramática que resulte de eliminar esa producción.

Se simplifica y se calcula la forma normal que nos pidan; tanto si es la FNCh, como si es la FNG, *después de obtenerla* se hace algo similar a lo indicado en la página 8 del boletín 5: es entonces cuando se introduce un nuevo símbolo inicial, S', cuyas únicas producciones serán

$$S' \to \lambda \mid S$$

siendo S el símbolo inicial de la gramática en forma normal que acabamos de obtener. La gramática que nos piden debe contener este nuevo inicial y estas dos producciones.

4.1.2. ¿Qué son todos esos enunciados que piden que calculemos un AP a partir de una gramática?

Bueno, yo les llamo FNG "enmascaradas" :-)

Sheyla Greibach demostró en su tesis doctoral que era posible construir autómatas de pila en los que en cada transición se consumiera un símbolo de la cinta de entrada. Si os fijáis, los únicos autómatas que definimos *intrínsecamente* como no deterministas son los APs, precisamente por esa propiedad de que es posible que se efectúen transiciones sin consumir símbolo de la cadena de entrada. Esto facilita la manipulación de la pila, por ejemplo.

Durante algún tiempo se discutió si era posible o no construir APs en los que cada transición consumiera un símbolo de la cinta de entrada. Sheyla Greibach demostró en su tesis que sí: de su Forma Normal se deriva una forma de construir autómatas de pila que aseguran este comportamiento.

Cada producción en FNG, es de la forma $A \to a\alpha$, con a un símbolo terminal y α una cadena Σ_A^*. La teoría nos dice (Tema 5):

Sea L un LCL, tal que $\lambda \notin L$, entonces existe una GCL $G = \langle \Sigma_A, \Sigma_T, P, S \rangle$ en Forma Normal de Greibach, tal que L(G) = L.

Se construye el siguiente AP, $A = \langle \Sigma, Q, \Gamma, q_0, Z_0, f, \emptyset \rangle$ donde,

$$\begin{array}{l} \Sigma = \Sigma_T \\ Q = \{q_0\} \\ \Gamma = \Sigma_A \\ Z_0 = S \\ (q_0, \alpha) \in f(q_0, a, A) \Leftrightarrow (A \to a\alpha) \in P \end{array}$$

y eso es lo que hacemos. Si volvemos al ejemplo,

$$\begin{array}{l} C_b \to b \\ C_a \to a \\ A \to a \mid aX \\ S \to aBC_a \mid aXBC_a \mid aC_a \mid aXC_a \mid a \mid aA \mid aXA \mid aAA \mid aXAA \\ B \to aBC_a \mid aXBC_a \mid aC_a \mid aXC_a \mid aC_b \mid aXC_b \\ X \to aBC_aC_a \mid aXBC_aC_a \mid aC_aC_a \mid aXC_aC_a \mid aC_bC_a \mid aXC_bC_a \mid a \mid \\ \qquad aBC_aC_aX \mid aXBC_aC_aX \mid aC_aC_aX \mid aXC_aC_aX \mid aC_bC_aX \mid aXC_bC_aX \mid aX \end{array}$$

el autómata asociado es: $A = \langle \Sigma, \Gamma, Q, f, q_0, Z_0, \emptyset \rangle$, con $\Sigma = \{a, b\}$, $\Gamma = \{C_b, C_a, A, S, B, X\}$, $Q = \{q_0\}$, $Z_0 = S$ y

$$\begin{array}{l} f(q_0, b, C_b) = \{(q_0, \lambda)\} \;\; /* \, Porque\ C_b \to b * / \\ f(q_0, a, C_a) = \{(q_0, \lambda)\} \;\; /* \, Porque\ C_a \to a * / \\ f(q_0, a, A) = \{(q_0, \lambda), (q_0, X)\} \;\; /* \, Porque\ A \to a \mid aX * / \\ f(q_0, a, S) = \{(q_0, BC_a), (q_0, XBC_a), (q_0, C_a), (q_0, XC_a), (q_0, \lambda), \\ \qquad\qquad (q_0, A), (q_0, XA), (q_0, AA), (q_0, XAA)\} \;\; /* \, Por\ qué?\ ; -) * / \\ f(q_0, a, B) = \{(q_0, BC_a), (q_0, XBC_a), (q_0, C_a), (q_0, XC_a), (q_0, C_b), (q_0, XC_b)\} \\ f(q_0, a, X) = \{(q_0, BC_aC_a), (q_0, XBC_aC_a), (q_0, C_aC_a), (q_0, XC_aC_a), (q_0, C_bC_a), \\ \qquad\qquad (q_0, XC_bC_a), (q_0, \lambda), (q_0, BC_aC_aX), (q_0, XBC_aC_aX)\}, (q_0, C_aC_aX), \\ \qquad\qquad (q_0, XC_aC_aX), (q_0, C_bC_aX), (q_0, XC_bC_aX), (q_0, X)\} \end{array}$$

Este autómata reconoce una cadena por pila vacía. Si se pidiera un AP por estado final se aplica el teorema 5.3:

Si $L = N(A)$ para algún AP A, entonces existe un AP A' tal que $L(A') = L$.

que proporciona un método para calcular un AP por estado final a partir de un AP por pila vacía. Si $A = \langle \Sigma, Q, \Gamma, f, q_0, Z_0, \emptyset \rangle$ es un AP por pila vacía, para construir un AP A' que reconozca L por estado final, lo definimos como $A' = \langle \Sigma, Q', \Gamma', f', q_0', X_0, F \rangle$ tal que

- $Q' = Q \cup \{q_0', q_f\},\ tal\ que\ q_0', q_f \notin Q$,
- $\Gamma' = \Gamma \cup \{X_0\},\ tal\ que\ X_0 \notin \Gamma$,
- $F = \{q_f\}$,

y en el que las reglas que definen a la función de transición f' son las siguientes:

1. $f'(q_0', \lambda, X_0) = \{(q_0, Z_0X_0)\}$,
2. $f'(q, a, B) = f(q, a, B),\ \forall\, q \in Q,\ \forall\, a \in (\Sigma \cup \{\lambda\}),\ \forall\, B \in \Gamma$,
3. $f'(q, \lambda, X_0) = \{(q_f', \lambda)\},\ \forall\, q \in Q$.

En el ejemplo, sería: $A' = \langle \Sigma, \Gamma', Q', f', q_0', X_0, F \rangle$, con $\Sigma = \{a, b\}$, $\Gamma' = \{C_b, C_a, A, S, B, X, X_0\}$, $Q = \{q_0, q_0', q_f\}$, y

$$
\begin{array}{l}
\mathbf{f'(q_0', \lambda, X_0) = \{(q_0, Z_0X_0)\}} \\
f'(q_0, b, C_b) = \{(q_0, \lambda)\} \\
f'(q_0, a, C_a) = \{(q_0, \lambda)\} \\
f'(q_0, a, A) = \{(q_0, \lambda), (q_0, X)\} \\
f'(q_0, a, S) = \{(q_0, BC_a), (q_0, XBC_a), (q_0, C_a), (q_0, XC_a), (q_0, \lambda), \\
\qquad\qquad\qquad (q_0, A), (q_0, XA), (q_0, AA), (q_0, XAA)\} \\
f'(q_0, a, B) = \{(q_0, BC_a), (q_0, XBC_a), (q_0, C_a), (q_0, XC_a), (q_0, C_b), (q_0, XC_b)\} \\
f'(q_0, a, X) = \{(q_0, BC_aC_a), (q_0, XBC_aC_a), (q_0, C_aC_a), (q_0, XC_aC_a), (q_0, C_bC_a), \\
\qquad\qquad\qquad (q_0, XC_bC_a), (q_0, \lambda), (q_0, BC_aC_aX), (q_0, XBC_aC_aX)\}, (q_0, C_aC_aX), \\
\qquad\qquad\qquad (q_0, XC_aC_aX), (q_0, C_bC_aX), (q_0, XC_bC_aX), (q_0, X)\} \\
\mathbf{f'(q_0, \lambda, X_0) = \{(q_f, \lambda)\}}
\end{array}
$$

5. Autoevaluación.

1. Indicar un posible orden de substitución en la siguiente gramática (en forma “pseudo-Chomsky”) de la que se pretende obtener la FNG:

$$
\begin{array}{l}
C_a \to a \\
C_b \to b \\
S \to C_aSC_b \mid C_aC_b \mid C_aAC_b \\
A \to BC_b \mid BAC_b \\
B \to SC_b \mid ASC_b \mid C_aC_bB
\end{array}
$$

<u>Solución:</u>

a) Me fijo en el primer símbolo del consecuente:

$$\begin{array}{l} C_a \rightarrow a \\ C_b \rightarrow b \\ S \rightarrow \mathbf{C_a} S C_b \mid \mathbf{C_a} C_b \mid \mathbf{C_a} A C_b \ (\Rightarrow C_a > S) \\ A \rightarrow \mathbf{B} C_b \mid \mathbf{B} A C_b \ (\Rightarrow B > A) \\ B \rightarrow \mathbf{S} C_b \mid \mathbf{A} S C_b \mid \mathbf{C_a} C_b B \ (\Rightarrow S > B, A > B, C_a > B) \end{array}$$

Y hay una contradicción: ¿$A > B$ y $B > A$? Opto por substituir los consecuentes de las producciones de A en las producciones de B.

b) Y me fijo otra vez en el primer símbolo del consecuente:

$$\begin{array}{l} C_a \rightarrow a \\ C_b \rightarrow b \\ S \rightarrow \mathbf{C_a} S C_b \mid \mathbf{C_a} C_b \mid \mathbf{C_a} A C_b \ (\Rightarrow C_a > S) \\ A \rightarrow \mathbf{B} C_b \mid \mathbf{B} A C_b \ (\Rightarrow B > A) \\ B \rightarrow \mathbf{S} C_b \mid \mathbf{B} C_b S C_b \mid \mathbf{B} A C_b S C_b \mid \mathbf{C_a} C_b B \ (\Rightarrow S > B, C_a > B \ y \ \ldots B > B??) \end{array}$$

Al substituir aparece una Recursividad a Izquierdas que debo eliminar.

c) Dicho y hecho, y vuelvo a fijarme en el primer símbolo del consecuente:

$$\begin{array}{l} C_a \rightarrow a \\ C_b \rightarrow b \\ S \rightarrow \mathbf{C_a} S C_b \mid \mathbf{C_a} C_b \mid \mathbf{C_a} A C_b \ (\Rightarrow C_a > S) \\ A \rightarrow \mathbf{B} C_b \mid \mathbf{B} A C_b \ (\Rightarrow B > A) \\ B \rightarrow \mathbf{S} C_b \mid \mathbf{C_a} C_b B \mid \mathbf{S} C_b X \mid \mathbf{C_a} C_b B X \ (\Rightarrow S > B, C_a > B) \\ X \rightarrow \mathbf{C_b} S C_b \mid \mathbf{A} C_b S C_b \mid \mathbf{C_b} S C_b X \mid \mathbf{A} C_b S C_b X \ (\Rightarrow C_b > X, A > X) \end{array}$$

El orden $C_a > C_b > S > B > A > X$ cumple las restricciones detectadas.

2. Dada la GCL G_1 cuyas producciones se exponen a continuación,

$$\begin{array}{l} S \rightarrow AB \mid ABC \mid a \\ A \rightarrow aA \mid aB \mid b \\ B \rightarrow BB \mid CA \\ C \rightarrow cS \mid cA \end{array}$$

a) Calcular una GCL equivalente a G_1 en forma normal de Greibach.

b) Calcular un AP que reconozca las mismas cadenas que las generadas por G_1.

<u>Solución:</u>

a) Gramática en FNG:

$$\begin{array}{l} C_c \rightarrow c \\ C_a \rightarrow a \\ A \rightarrow aA \mid aB \mid b \\ S \rightarrow aAB \mid aBB \mid bB \mid aABC \mid aBBC \mid bBC \mid a \\ C \rightarrow cS \mid cA \\ B \rightarrow cSA \mid cAA \mid cSAX \mid cAAX \\ X \rightarrow cSA \mid cAA \mid cSAX \mid cAAX \mid cSAX \mid cAAX \mid cSAXX \mid cAAXX \end{array}$$

b) AP por pila vacía: $A = \langle \Sigma, \Gamma, Q, f, q_0, Z_0, \emptyset \rangle$, con $\Sigma = \{a, b, c\}$, $\Gamma = \{C_c, C_a, S, A, B, C, X\}$, $Q = \{q_0\}$, $Z_0 = S$ y

$$
\begin{array}{l}
f(q_0, c, C_c) = \{(q_0, \lambda)\} \\
f(q_0, a, C_a) = \{(q_0, \lambda)\} \\
f(q_0, a, A) = \{(q_0, A), (q_0, B)\} \\
f(q_0, b, A) = \{(q_0, \lambda)\} \\
f(q_0, a, S) = \{(q_0, AB), (q_0, BB), (q_0, ABC), (q_0, BBC), (q_0, \lambda)\} \\
f(q_0, b, S) = \{(q_0, B), (q_0, BC)\} \\
f(q_0, c, C) = \{(q_0, S), (q_0, A)\} \\
f(q_0, c, B) = \{(q_0, SA), (q_0, AA), (q_0, SAX), (q_0, AAX)\} \\
f(q_0, c, X) = \{(q_0, SA), (q_0, AA), (q_0, SAX), (q_0, AAX), (q_0, SAX), (q_0, AAX), \\
\qquad\qquad\qquad (q_0, SAXX), (q_0, AAXX)\}
\end{array}
$$

3. Sea G_1 la GCL definida por el siguiente conjunto de producciones:

$$
\begin{array}{l}
S \to CBC \mid CA \mid a \\
A \to SBC \mid b \\
B \to BB \mid b \\
C \to aa
\end{array}
$$

a) Escribir una GCL, equivalente a G_1, en Forma Normal de Greibach.

b) Escribir un AP por pila vacía y otro por estado final que reconozcan $L(G_1)$.

<u>Solución:</u>

a) Gramática en FNG:

$$
\begin{array}{l}
C_a \to a \\
C \to aC_a \\
S \to aC_aBC \mid aC_aA \mid a \\
A \to aC_aBCBC \mid aC_aABC \mid aBC \mid b \\
B \to b \mid bX \\
X \to b \mid bX \mid bXX
\end{array}
$$

b) 1) AP por pila vacía: $A = \langle \Sigma, \Gamma, Q, f, q_0, Z_0, \emptyset \rangle$, con $\Sigma = \{a, b\}$, $\Gamma = \{C_a, C, S, A, B, X\}$, $Q = \{q_0\}$, $Z_0 = S$ y

$$
\begin{array}{l}
f(q_0, a, C_a) = \{(q_0, \lambda)\} \\
f(q_0, a, C) = \{(q_0, C_a)\} \\
f(q_0, a, S) = \{(q_0, C_aBC), (q_0, C_aA), (q_0, \lambda)\} \\
f(q_0, a, A) = \{(q_0, C_aBCBC), (q_0, C_aABC), (q_0, BC)\} \\
f(q_0, b, A) = \{(q_0, \lambda)\} \\
f(q_0, b, B) = \{(q_0, \lambda), (q_0, X)\} \\
f(q_0, b, X) = \{(q_0, \lambda), (q_0, X), (q_0, XX)\}
\end{array}
$$

2) AP por estado final: $A' = \langle \Sigma, \Gamma', Q', f', q_0', X_0, F \rangle$, con $\Sigma = \{a, b\}$, $\Gamma' = \Gamma \cup \{X_0\}$,

$Q' = Q \cup \{q_0', q_f\}$, $F = \{q_f\}$ y

$$
\begin{array}{l}
\mathbf{f'(q_0', \lambda, X_0) = \{(q_0, Z_0X_0)\}} \\
f(q_0, a, C_a) = \{(q_0, \lambda)\} \\
f(q_0, a, C) = \{(q_0, C_a)\} \\
f(q_0, a, S) = \{(q_0, C_aBC), (q_0, C_aA), (q_0, \lambda)\} \\
f(q_0, a, A) = \{(q_0, C_aBCBC), (q_0, C_aABC), (q_0, BC)\} \\
f(q_0, b, A) = \{(q_0, \lambda)\} \\
f(q_0, b, B) = \{(q_0, \lambda), (q_0, X)\} \\
f(q_0, b, X) = \{(q_0, \lambda), (q_0, X), (q_0, XX)\} \\
\mathbf{f'(q_0, \lambda, X_0) = \{(q_f, \lambda)\}}
\end{array}
$$

4. Calcular un AP por estado final y otro por pila vacía que reconozca el mismo lenguaje que genera la siguiente GCL:

$$
\begin{array}{l}
S \to aS \mid bA \mid ba \\
A \to ASa \mid bSA \mid cBB \mid a \\
B \to bBS \mid aABS \mid cBS
\end{array}
$$

Solución:

a) AP por pila vacía: $A = \langle \Sigma, \Gamma, Q, f, q_0, Z_0, \emptyset \rangle$, con $\Sigma = \{a, b\}$, $\Gamma = \{C_a, C_b, S, A, X\}$, $Q = \{q_0\}$, $Z_0 = S$ y

$$
\begin{array}{l}
f(q_0, a, C_a) = \{(q_0, \lambda)\} \\
f(q_0, b, C_b) = \{(q_0, \lambda)\} \\
f(q_0, a, S) = \{(q_0, S)\} \\
f(q_0, b, S) = \{(q_0, a), (q_0, C_a)\} \\
f(q_0, a, A) = \{(q_0, \lambda), (q_0, X)\} \\
f(q_0, b, A) = \{(q_0, SA), (q_0, SAX)\} \\
f(q_0, a, X) = \{(q_0, SC_a), (q_0, SC_aX)\} \\
f(q_0, b, X) = \{(q_0, AC_a), (q_0, C_aC_a), (q_0, AC_aX), (q_0, C_aC_aX)\}
\end{array}
$$

b) AP por estado final: $A' = \langle \Sigma, \Gamma', Q', f', q_0', X_0, F \rangle$, con $\Sigma = \{a, b\}$, $\Gamma' = \Gamma \cup \{X_0\}$, $Q' = Q \cup \{q_0', q_f\}$, $F = \{q_f\}$ y

$$
\begin{array}{l}
\mathbf{f'(q_0', \lambda, X_0) = \{(q_0, Z_0X_0)\}} \\
f'(q_0, a, C_a) = \{(q_0, \lambda)\} \\
f'(q_0, b, C_b) = \{(q_0, \lambda)\} \\
f'(q_0, a, S) = \{(q_0, S)\} \\
f'(q_0, b, S) = \{(q_0, a), (q_0, C_a)\} \\
f'(q_0, a, A) = \{(q_0, \lambda), (q_0, X)\} \\
f'(q_0, b, A) = \{(q_0, SA), (q_0, SAX)\} \\
f'(q_0, a, X) = \{(q_0, SC_a), (q_0, SC_aX)\} \\
f'(q_0, b, X) = \{(q_0, AC_a), (q_0, C_aC_a), (q_0, AC_aX), (q_0, C_aC_aX)\} \\
\mathbf{f'(q_0, \lambda, X_0) = \{(q_f, \lambda)\}}
\end{array}
$$

5. Dada la GCL,

$$
\begin{array}{l}
S \to SA \mid a \\
A \to SB \mid SS \mid b \\
B \to AS \mid SB
\end{array}
$$

a) Calcular una GCL equivalente en forma normal de Greibach.

b) Calcular un AP que reconozca las mismas cadenas que las generadas por la gramática.

<u>Solución:</u>

a)

$$
\begin{aligned}
&S \to a \mid aX \\
&A \to aB \mid aXB \mid aS \mid aXS \mid b \\
&B \to aBS \mid aXBS \mid aSS \mid aXSS \mid bS \mid aB \mid aXB \\
&X \to aB \mid aXB \mid aS \mid aXS \mid b \mid aBX \mid aXBX \mid aSX \mid aXSX \mid bX
\end{aligned}
$$

b) AP por pila vacía: $A = \langle \Sigma, \Gamma, Q, f, q_0, Z_0, \emptyset \rangle$, con $\Sigma = \{a, b\}$, $\Gamma = \{S, A, B, X\}$, $Q = \{q_0\}$, $Z_0 = S$ y

$$
\begin{aligned}
&f(q_0, a, S) = \{(q_0, \lambda), (q_0, X)\} \\
&f(q_0, a, A) = \{(q_0, B), (q_0, XB), (q_0, S), (q_0, XS)\} \\
&f(q_0, b, A) = \{(q_0, \lambda)\} \\
&f(q_0, a, B) = \{(q_0, BS), (q_0, XBS), (q_0, SS), (q_0, XSS), (q_0, B), (q_0, XB)\} \\
&f(q_0, b, B) = \{(q_0, S)\} \\
&f(q_0, a, X) = \{(q_0, B), (q_0, XB), (q_0, S), (q_0, XS), (q_0, BX), (q_0, XBX), (q_0, SX), (q_0, XSX)\} \\
&f(q_0, b, X) = \{(q_0, \lambda), (q_0, X)\}
\end{aligned}
$$