

## Ejercicio Guía para obtener FNC y FNG

Este documento tiene como fin mostrar la forma de obtener la Forma Normal de Chomsky (FNC o también CNF) y la Forma Normal de Greibach (FNG o también GNF) de una gramática dada, mediante un ejercicio explicativo.

La gramática en cuestión del ejercicio es la siguiente:

$S \rightarrow A \mid BCa \mid aDcd \mid EDF$   
 $A \rightarrow aAb \mid c$   
 $B \rightarrow CD \mid b \mid ECd \mid Ad$   
 $C \rightarrow Cc \mid Bb \mid AaE \mid \epsilon$   
 $D \rightarrow aDd \mid Dd \mid \epsilon$   
 $E \rightarrow aaEB \mid EFG$   
 $F \rightarrow aFd \mid d$

Durante este ejercicio, la gramática irá sufriendo modificaciones que en algunos casos mantendrá a su lenguaje intacto y en otras lo alterará.

### Forma Intermedia de Chomsky

Antes de empezar con el proceso para obtener la FNC de una gramática, debemos dejar a dicha gramática en un estado intermedio, al cual podemos llamar forma intermedia de Chomsky, aunque no suele llamársele por ningún nombre. De todas formas, en este paso, lo que vamos a tratar de obtener es una gramática libre de símbolos inútiles, producciones  $\epsilon$  y producciones unitarias. Cuando ya hayamos limpiado a nuestra gramática de estas cosas, podremos empezar a obtener su FNC.

### *Eliminando Símbolos Inútiles*

En este paso, eliminaremos todos aquellos símbolos que sean inútiles. Hay dos tipos de símbolos inútiles: los no generadores y los no alcanzables.

### **Eliminando Símbolos No Generadores**

Para eliminar símbolos no generadores, primero tratamos de identificar todos aquellos símbolos que sean generadores. Aquellas variables que no podamos determinar que son generadoras pasarán a ser símbolos no generadores y se eliminarán.

Por definición, todos los símbolos terminales de una gramática son generadores, ya que se generan a sí mismos.

También será un símbolo generador toda aquella variable que tenga por lo menos una producción que esté conformada únicamente de símbolos generadores. El string vacío  $\epsilon$  se considera un símbolo generador.

Identifiquemos en nuestra gramática nuestros símbolos generadores: A es generador, ya que tiene la producción  $A \rightarrow c$ . También lo es B, ya que tiene  $B \rightarrow b$ . C y D son generadores, ya que tienen la producción vacía  $C \rightarrow \epsilon$  y  $D \rightarrow \epsilon$ , respectivamente. F tiene la producción  $F \rightarrow d$ . Por último, S tiene la producción  $S \rightarrow BCa$ , donde tanto B, C como a son símbolos generadores, por lo tanto S es generador también.

E no es un símbolo generador. Tiene dos producciones, es verdad. Una es  $E \rightarrow aaEB$  y la otra es  $E \rightarrow EFG$ , y en las dos hay presentes símbolos generadores, como por ejemplo a y B en la primera producción y F en la segunda. Pero ninguna de dichas producciones está completamente compuesta por símbolos generadores. Miremos por ejemplo la segunda producción: está presente la variable G, que no tiene ninguna producción asociada, por lo que es incapaz de ser un símbolo generador. Descartamos esta producción entonces. Veamos la primera producción, tal vez descubramos algo: Mmmmh, tiene a la variable

E entre medio, la cual todavía no sabemos si es generadora. Qué recursivo, ¿eh? Pues bueno, la recursión llega hasta ahí: No sabemos si E es generador y no hay ninguna otra producción de E que nos ayude a verificar esto, así que la primera producción no nos sirve para determinar a E como un símbolo generador, y como no podemos verificar que E es un símbolo generador, entonces lo consideramos como un símbolo NO generador. Procedemos ahora a eliminar las variables E y G de la gramática y a todas las producciones en las que estén presentes. O sea, eliminamos  $S \rightarrow EDF$ ,  $B \rightarrow ECd$ ,  $C \rightarrow AaE$ ,  $E \rightarrow aaEB$  y  $E \rightarrow EF$ . La gramática queda como sigue:

$S \rightarrow A \mid BCa \mid aDcd$   
 $A \rightarrow aAb \mid c$   
 $B \rightarrow CD \mid b \mid Ad$   
 $C \rightarrow Cc \mid Bb \mid \epsilon$   
 $D \rightarrow aDd \mid Dd \mid \epsilon$   
 $F \rightarrow aFd \mid d$

Puede que al lector le importe saber que no se altera el lenguaje de la gramática cuando se eliminan los símbolos no generadores.

### Eliminando Símbolos No Alcanzables

Puede que, desde un principio o por producto de la eliminación de algunos símbolos no generadores, nuestra gramática presente símbolos no alcanzables o, dicho de otras palabras, símbolos que no son derivables (directa o indirectamente) desde el símbolo de inicio S. En nuestro ejemplo particular, nuestra querida variable F sufre de esta injusticia: ya no existe producción que la derive desde el símbolo inicial S, pues dichas producciones se eliminaron cuando estábamos eliminando las variables no generadoras. ¿Qué hacemos entonces? Somos más injustos aún y la sacamos del grupo también. La gramática nos queda algo así:

$S \rightarrow A \mid BCa \mid aDcd$   
 $A \rightarrow aAb \mid c$   
 $B \rightarrow CD \mid b \mid Ad$   
 $C \rightarrow Cc \mid Bb \mid \epsilon$   
 $D \rightarrow aDd \mid Dd \mid \epsilon$

La gramática nos está quedando cada vez menos poblada, pero no se preocupen que pronto las producciones y las variables van a proliferar en nuestra gramática como conejitos jóvenes y enérgicos.

### Eliminando Producciones $\epsilon$

Ahora viene un paso un poco más difícil. La eliminación de producciones  $\epsilon$  consiste principalmente en eliminar las producciones de la forma  $X \rightarrow \epsilon$ . Cuando uno haga esto, lo que tendrá que hacer es contemplar la posibilidad de que, en donde antes aparecía la variable X, ésta ahora podrá aparecer o no aparecer, ya que en una estará jugando el papel de ser el string vacío y en la otra se estará siendo otra cosa.

¡Ah! ¡Ya estamos listos entonces! Bueno, no del todo: ¿Qué pasa si una variable tiene una producción que está compuesta de varios símbolos que pueden ser reemplazables por el string vacío? Por ejemplo, digamos que tenemos  $Y \rightarrow X_1X_2X_3\dots X_n$ , donde todos los  $X_i$  pueden ser el string  $\epsilon$ . Entonces nos quedaría una cosa así:  $Y \rightarrow \epsilon\epsilon\dots\epsilon$ , que se puede simplificar por  $Y \rightarrow \epsilon$ . Aquí vemos que Y también puede generar al string vacío. Entonces vamos a definir a los símbolos **Anulables**. Un símbolo anulable Y va a ser aquel que tenga entre sus producciones a la producción  $Y \rightarrow \epsilon$  o también aquel que tenga por lo menos una producción en donde todas las variables involucradas son anulables a su vez.

Viendo nuestro ejemplo, podemos identificar rápidamente a dos variables que son símbolos anulables: las variables C y D, ya que ambas tienen la producción de la forma  $Y \rightarrow \epsilon$ . Pero ojo, que no son las únicas variables anulables. También está B que tiene la producción  $B \rightarrow CD$ , la cual está compuesta enteramente por símbolos anulables.

Bien, procedamos a eliminar las producciones  $\epsilon$ . ¿Cómo hacemos esto? Muy fácil: en cada producción en la que estaban presentes uno o más símbolos anulables, vemos todas las combinaciones en las que cada una de dichas variables pudiera estar presente o no. Por ejemplo, si tenemos la producción

$Y \rightarrow abX_1$ , donde  $X_1$  es anulable, entonces se generan las producciones  $Y \rightarrow abX_1$  y  $Y \rightarrow ab$ , una en donde  $X_1$  está presente y otra en la que no. Veamos el caso para producciones con dos y tres variables anulables:

$Y \rightarrow aX_1bbX_2 \Rightarrow Y \rightarrow aX_1bbX_2 \mid aX_1bb \mid abbX_2 \mid abb$

$Y \rightarrow cX_1X_2ddX_3a \Rightarrow Y \rightarrow cX_1X_2ddX_3a \mid cX_1X_2dda \mid cX_1ddX_3a \mid cX_2ddX_3a \mid cX_1dda \mid cX_2dda \mid cddX_3a \mid cdda$

En general, si tenemos una producción con  $n$  símbolos anulables, tendremos que reemplazar esa producción con a lo más  $2^n$  producciones. En nuestro caso, la gramática quedaría como sigue:

$S \rightarrow A \mid BCa \mid Ba \mid Ca \mid a \mid aDcd \mid acd$

$A \rightarrow aAb \mid c$

$B \rightarrow CD \mid C \mid D \mid b \mid Ad$

$C \rightarrow Cc \mid c \mid Bb \mid b$

$D \rightarrow aDd \mid ad \mid Dd \mid d$

Miren como de repente ha crecido nuestra querida gramática. No han aparecido nuevas variables, pero sí surgieron más producciones. Pero esto no se detiene aquí. Ojo, que ahora sí que se puede haber alterado el lenguaje. Si nuestra gramática inicial era capaz de generar el string vacío, ahora no lo podrá generar.

## ***Eliminando Producciones Unitarias***

Las producciones unitarias son aquellas que tienen una única variable en su cuerpo (ojo!! variable!!! símbolos terminales solos no cuentan como producciones unitarias!!). Si tenemos por ejemplo una producción  $Y \rightarrow X$ , la idea es reemplazar esta producción por todas las producciones de  $X$ . ¿Pero qué pasa si  $X$  también tiene producciones unitarias? Bueno, entonces primero reemplazamos esas producciones unitarias por sus correspondientes producciones. ¿Y si  $X$  tiene la producción unitaria  $X \rightarrow Y$ ? Entonces ahí lo que hacemos es primero poner en  $Y$  todas las producciones de  $X$  menos la producción  $X \rightarrow Y$ , y luego ponemos en  $X$  todas las producciones de  $Y$ , menos la producción  $Y \rightarrow X$ .

Hay una forma más general de describir este proceso, que involucra la definición de los pares unitarios  $(Y,X)$ , que se definen así:

**Base:** El par  $(Y,Y)$  es unitario por definición, para toda variable de una gramática dada.

**Inducción:** Si el par  $(Y,X)$  es unitario y la variable  $X$  tiene una producción unitaria de la forma  $X \rightarrow Z$ , entonces  $(Y,Z)$  también es un par unitario.

Nótese que los pares unitarios  $(X,Y)$  e  $(Y,X)$  son distintos. Lo que se hace a continuación es que por cada par unitario  $(X,Y)$ , agregamos a las producciones de  $X$  todas aquellas producciones de  $Y$  que NO son unitarias. En nuestra gramática tenemos los siguientes pares unitarios:  $(S,S)$ ,  $(S,A)$ ,  $(A,A)$ ,  $(B,B)$ ,  $(B,C)$ ,  $(B,D)$ ,  $(C,C)$  y  $(D,D)$ . Los más relevantes de estos pares son  $(S,A)$ ,  $(B,C)$  y  $(B,D)$ . Procedemos a eliminar las producciones unitarias:

$S \rightarrow aAb \mid c \mid BCa \mid Ba \mid Ca \mid a \mid aDcd \mid acd$

$A \rightarrow aAb \mid c$

$B \rightarrow CD \mid Cc \mid c \mid Bb \mid b \mid aDd \mid ad \mid Dd \mid d \mid Ad$

$C \rightarrow Cc \mid c \mid Bb \mid b$

$D \rightarrow aDd \mid ad \mid Dd \mid d$

Este tipo de modificaciones tampoco altera el lenguaje de la gramática.

## **Forma Normal de Chomsky**

Ya hemos dejado a nuestra gramática inicial en una forma intermedia apta para ser modificada en una FNC. Lo que hacemos a continuación son dos pasos:

### ***Reemplazar Terminales por Variables***

El primer paso consiste en reemplazar todos aquellos símbolos terminales que están acompañados en alguna producción por uno o más símbolos. A estos terminales los reemplazamos por nuevas variables

que tendrán una única producción (la producción de dicho símbolo). En nuestro caso, hay que crear 4 nuevas variables, una para cada símbolo terminal:

$$\begin{aligned} S &\rightarrow A_1AB_1 \mid c \mid BCA_1 \mid BA_1 \mid CA_1 \mid a \mid A_1DC_1D_1 \mid A_1C_1D_1 \\ A &\rightarrow A_1AB_1 \mid c \\ B &\rightarrow CD \mid CC_1 \mid c \mid BB_1 \mid b \mid A_1DD_1 \mid A_1D_1 \mid DD_1 \mid d \mid AD_1 \\ C &\rightarrow CC_1 \mid c \mid BB_1 \mid b \\ D &\rightarrow A_1DD_1 \mid A_1D_1 \mid DD_1 \mid d \\ A_1 &\rightarrow a \\ B_1 &\rightarrow b \\ C_1 &\rightarrow c \\ D_1 &\rightarrow d \end{aligned}$$

Este tipo de modificaciones no altera el lenguaje de la gramática.

## Reemplazar Producciones de 3 o más Variables

El segundo paso consiste en eliminar todas las producciones con 3 o más variables por una que contenga sólo dos. ¿Cómo hacemos esto? Supongamos que tenemos la producción  $Y \rightarrow X\alpha$ , donde  $\alpha$  es un string que consta de dos o más variables. Entonces creamos una nueva variable  $Z$ , con la producción  $Z \rightarrow \alpha$ , y luego reemplazamos  $Y \rightarrow X\alpha$  por  $Y \rightarrow XZ$ . ¿Y si  $\alpha$  tiene más de dos variables? Supongamos que  $\alpha$  está compuesto por la variable  $W$  al principio y por  $\beta$  al final, donde  $\beta$  es a su vez un string que tiene dos o más variables, entonces creamos una variable  $V$  con la producción  $V \rightarrow \beta$  y luego reemplazamos  $Z \rightarrow W\alpha$  por  $Z \rightarrow WV$ . Repetimos este proceso hasta que ya no queden producciones de 3 o más variables en nuestra gramática. En nuestro caso, la gramática queda como sigue:

$$\begin{aligned} S &\rightarrow A_1Y_1 \mid c \mid BY_2 \mid BA_1 \mid CA_1 \mid a \mid A_1Y_3 \mid A_1Y_4 \\ Y_1 &\rightarrow AB_1 \\ Y_2 &\rightarrow CA_1 \\ Y_3 &\rightarrow DY_4 \\ Y_4 &\rightarrow C_1D_1 \\ A &\rightarrow A_1Y_1 \mid c \\ B &\rightarrow CD \mid CC_1 \mid c \mid BB_1 \mid b \mid A_1Y_5 \mid A_1D_1 \mid DD_1 \mid d \mid AD_1 \\ Y_5 &\rightarrow DD_1 \\ C &\rightarrow CC_1 \mid c \mid BB_1 \mid b \\ D &\rightarrow A_1Y_5 \mid A_1D_1 \mid DD_1 \mid d \\ A_1 &\rightarrow a \\ B_1 &\rightarrow b \\ C_1 &\rightarrow c \\ D_1 &\rightarrow d \end{aligned}$$

Este tipo de modificaciones no altera el lenguaje de la gramática. Acabamos de obtener una gramática que tiene producciones que involucran o solamente a dos variables ( $X \rightarrow YZ$ ) o a un único símbolo terminal ( $W \rightarrow v$ ). Este tipo de gramáticas, damas y caballeros, se denominan gramáticas en la Forma Normal de Chomsky, y tienen la particularidad de que todos los árboles de derivación que generan son completamente binarios. Esta característica los hace bastante útiles para el problema de identificar si un string dado  $w$  pertenece al lenguaje de dicha gramática.

## Forma Intermedia de Greibach

Al igual que con la FNC, una gramática dada debe ser llevada a una forma intermedia antes de poder transformarla a su correspondiente forma normal de Greibach. En el caso de Chomsky, la forma intermedia implicaba eliminar los símbolos inútiles, las producciones  $\epsilon$  y las producciones unitarias. En el caso de Greibach, la forma intermedia debe cumplir con que todas las producciones sean o strings de una o más variables, o producciones de un solo símbolo terminal, o producciones de un símbolo terminal seguido por una o más variables. Existen muchos métodos para llegar a esta forma intermedia. La que recomienda la mayoría de la gente es la que consiste en obtener la FNC de una gramática. ¿Y por qué? Porque una gramática que esté en FNC cumple con estas tres condiciones de la forma intermedia de Greibach. Realmente, cuando se han reemplazado a los símbolos terminales por variables en el proceso de formalización de Chomsky, ya en ese entonces tenemos una forma intermedia de Greibach. Uno podría elegir también esta gramática como su forma intermedia de Greibach.

## Forma Normal de Greibach

La forma normal de Greibach de una gramática dada es aquella que sólo contiene producciones de la forma  $Y \rightarrow x\alpha$ , donde  $x$  es un terminal y  $\alpha$  es un string conformado únicamente por variables y puede tener largo cero (o sea, puede ser el string vacío). Una vez tenemos nuestra gramática en su forma intermedia, procedemos a aplicarle las siguientes modificaciones, para así obtener su FNG:

1. Ordenamos y enumeramos nuestras variables de tal forma que nos quede nuestro símbolo inicial  $S$  como la primera variable de la enumeración. La gramática nos quedaría así:

$S$	$X_1 \rightarrow X_{11}X_2 \mid c \mid X_8X_3 \mid X_6X_{11} \mid X_8X_{11} \mid a \mid X_{11}X_4 \mid X_{11}X_5$
$Y_1$	$X_2 \rightarrow X_{10}X_{12}$
$Y_2$	$X_3 \rightarrow X_8X_{11}$
$Y_3$	$X_4 \rightarrow X_9X_5$
$Y_4$	$X_5 \rightarrow X_{13}X_{14}$
$B$	$X_6 \rightarrow X_8X_9 \mid X_8X_{13} \mid c \mid X_6X_{12} \mid b \mid X_{11}X_7 \mid X_{11}X_{14} \mid X_9X_{14} \mid d \mid X_{10}X_{14}$
$Y_5$	$X_7 \rightarrow X_9X_{12}$
$C$	$X_8 \rightarrow X_8X_{13} \mid c \mid X_8X_{12} \mid b$
$D$	$X_9 \rightarrow X_{11}X_7 \mid X_{11}X_{14} \mid X_9X_{14} \mid d$
$A$	$X_{10} \rightarrow X_{11}X_2 \mid c$
$A_1$	$X_{11} \rightarrow a$
$B_1$	$X_{12} \rightarrow b$
$C_1$	$X_{13} \rightarrow c$
$D_1$	$X_{14} \rightarrow d$

Dejamos el símbolo original a la izquierda de cada línea para que se entienda cuál fue la variable enumerada que reemplazó a cada variable. Nótese que cambiamos de orden a la variable  $A$ , al ponerla por debajo de la variable  $D$ . El propósito de estas alteraciones en el orden original tienen como fin que, al proceder con la enumeración, en las producciones resultantes quede la menor cantidad de producciones  $X_i \rightarrow X_j\alpha$ , en donde  $i$  sea mayor que  $j$ .

2. Eliminamos la recursión inmediata a la izquierda y la recursión a la izquierda. ¿Qué se entiende por recursión inmediata a la izquierda? Cuando tenemos una producción de la forma  $X \rightarrow X\alpha$ , decimos que  $X$  tiene recursividad inmediata por la izquierda. Cuando tenemos por lo menos una producción de la forma  $X \rightarrow Y\alpha$ , y la variable  $Y$  tiene por lo menos una producción de la forma  $Y \rightarrow X\beta$  decimos que  $X$  recursividad por la izquierda. ¿Cómo eliminamos la recursión por la izquierda? Primero reemplazamos en las producciones de  $X$  todas las apariciones de  $Y$  por aquellas producciones de  $Y$  que no sean inmediatamente recursivas a la izquierda para  $Y$ . Con esto logramos hacer que  $X$  tenga sólo producciones inmediatamente recursivas por la izquierda. Luego, supongamos que  $X$  tiene las siguientes producciones:

$$X \rightarrow X\alpha_1 \mid X\alpha_2 \mid \dots \mid X\alpha_n \mid \beta_1 \mid \beta_2 \mid \dots \mid \beta_m$$

Donde las producciones  $\beta_i$  no son recursivas por la izquierda. Ahora creamos una variable auxiliar  $X'$  y realizamos las siguientes transformaciones:

$$\begin{aligned} X &\rightarrow \beta_1 \mid \beta_2 \mid \dots \mid \beta_m \mid \beta_1X' \mid \beta_2X' \mid \dots \mid \beta_mX' \\ X' &\rightarrow \alpha_1 \mid \alpha_2 \mid \dots \mid \alpha_nX' \mid \alpha_1X' \mid \alpha_2X' \mid \dots \mid \alpha_nX' \end{aligned}$$

En nuestro caso, las variables que tienen producciones inmediatamente recursivas por la izquierda son  $X_6$ ,  $X_8$  y  $X_9$ , debido a sus producciones  $X_6 \rightarrow X_6X_{12}$ ,  $X_8 \rightarrow X_8X_{13}$  y  $X_9 \rightarrow X_9X_{14}$ . También tenemos la producción  $X_9 \rightarrow X_6X_{12}$  que es recursiva a la izquierda para la variable  $X_6$ , y las producciones  $X_6 \rightarrow X_8X_9$  y  $X_6 \rightarrow X_8X_{13}$  que son recursivas a la izquierda para la variable  $X_8$ . Procedemos a dejar a ambas variables únicamente con recursividad inmediata a la izquierda:

- 1) Reemplazando a  $X_8$  por sus producciones no recursivas a la izquierda, las producciones  $X_6 \rightarrow X_8X_3 \mid X_8X_{13} \mid c \mid X_6X_{12} \mid b \mid X_{11}X_7 \mid X_{11}X_{14} \mid X_9X_{14} \mid d \mid X_{10}X_{14}$  se transforman en:

$$X_6 \rightarrow cX_9 \mid X_6X_{12}X_9 \mid bX_9 \mid cX_{13} \mid X_6X_{12}X_{13} \mid bX_{13} \mid c \mid X_6X_{12} \mid b \mid X_{11}X_7 \mid X_{11}X_{14} \mid X_9X_{14} \mid d \mid X_{10}X_{14}$$

Nótese que no reemplazamos a  $X_8$  por su producción  $X_8 \rightarrow X_8X_{13}$ , ya que esta producción es recursiva por la izquierda para  $X_8$ .

- 2) Reemplazando a  $X_6$  por sus producciones no recursivas a la izquierda, las producciones  $X_0 \rightarrow X_0X_{13} \mid c \mid X_6X_{12} \mid b$  se transforman en:

$$X_8 \rightarrow X_8X_{13} \mid c \mid X_8X_9X_{12} \mid X_8X_{13}X_{12} \mid cX_{12} \mid bX_{12} \mid X_{11}X_7X_{12} \mid X_{11}X_{14}X_{12} \mid X_9X_{14}X_{12} \mid dX_{12} \mid X_{10}X_{14}X_{12} \mid b$$

Aquí tampoco reemplazamos a  $X_6$  por su producción  $X_6 \rightarrow X_6X_{12}$ , ya que esta producción es recursiva por la izquierda para  $X_6$ . Ahora que solamente tenemos recursividad inmediata por la izquierda, procedemos a realizar las correspondientes transformaciones:

- 1) Las producciones  $X_6 \rightarrow cX_9 \mid X_6X_{12}X_9 \mid bX_9 \mid cX_{13} \mid X_6X_{12}X_{13} \mid bX_{13} \mid c \mid X_6X_{12} \mid b \mid X_{11}X_7 \mid X_{11}X_{14} \mid X_9X_{14} \mid d \mid X_{10}X_{14}$  se transforman en:

$$\begin{aligned} X_6 &\rightarrow cX_9 \mid bX_9 \mid cX_{13} \mid bX_{13} \mid c \mid b \mid X_{11}X_7 \mid X_{11}X_{14} \mid X_9X_{14} \mid d \mid X_{10}X_{14} \mid cX_9X_{6,1} \mid bX_9X_{6,1} \mid cX_{13}X_{6,1} \mid bX_{13}X_{6,1} \mid \\ &\quad cX_{6,1} \mid bX_{6,1} \mid X_{11}X_7X_{6,1} \mid X_{11}X_{14}X_{6,1} \mid X_9X_{14}X_{6,1} \mid dX_{6,1} \mid X_{10}X_{14}X_{6,1} \\ X_{6,1} &\rightarrow X_{12}X_9 \mid X_{12}X_{13} \mid X_{12} \mid X_{12}X_9X_{6,1} \mid X_{12}X_{13}X_{6,1} \mid X_{12}X_{6,1} \end{aligned}$$

- 2) Las producciones  $X_8 \rightarrow X_8X_{13} \mid c \mid X_8X_9X_{12} \mid X_8X_{13}X_{12} \mid cX_{12} \mid bX_{12} \mid X_{11}X_7X_{12} \mid X_{11}X_{14}X_{12} \mid X_9X_{14}X_{12} \mid dX_{12} \mid X_{10}X_{14}X_{12} \mid b$  se transforman en:

$$\begin{aligned} X_8 &\rightarrow c \mid cX_{12} \mid bX_{12} \mid X_{11}X_7X_{12} \mid X_{11}X_{14}X_{12} \mid X_9X_{14}X_{12} \mid dX_{12} \mid X_{10}X_{14}X_{12} \mid b \mid cX_{8,1} \mid cX_{12}X_{8,1} \mid bX_{12}X_{8,1} \mid \\ &\quad X_{11}X_7X_{12}X_{8,1} \mid X_{11}X_{14}X_{12}X_{8,1} \mid X_9X_{14}X_{12}X_{8,1} \mid dX_{12}X_{8,1} \mid X_{10}X_{14}X_{12}X_{8,1} \mid bX_{8,1} \\ X_{8,1} &\rightarrow X_{13} \mid X_9X_{12} \mid X_{13}X_{12} \mid X_{13}X_{8,1} \mid X_9X_{12}X_{8,1} \mid X_{13}X_{12}X_{8,1} \end{aligned}$$

- 3) Y las producciones  $X_c \rightarrow X_{11}X_7 \mid X_{11}X_{14} \mid X_9X_{14} \mid d$  se transforman en:

$$\begin{aligned} X_9 &\rightarrow X_{11}X_7 \mid X_{11}X_{14} \mid d \mid X_{11}X_7X_{9,1} \mid X_{11}X_{14}X_{9,1} \mid dX_{9,1} \\ X_{9,1} &\rightarrow X_{14} \mid X_{14}X_{9,1} \end{aligned}$$

Nuestra gramática, libre de recursividad por la izquierda, queda como sigue:

$$\begin{aligned} X_1 &\rightarrow X_{11}X_2 \mid c \mid X_6X_3 \mid X_6X_{11} \mid X_8X_{11} \mid a \mid X_{11}X_4 \mid X_{11}X_5 \\ X_2 &\rightarrow X_{10}X_{12} \\ X_3 &\rightarrow X_6X_{11} \\ X_4 &\rightarrow X_9X_5 \\ X_5 &\rightarrow X_{13}X_{14} \\ X_6 &\rightarrow cX_9 \mid bX_9 \mid cX_{13} \mid bX_{13} \mid c \mid b \mid X_{11}X_7 \mid X_{11}X_{14} \mid X_9X_{14} \mid d \mid X_{10}X_{14} \mid cX_9X_{6,1} \mid bX_9X_{6,1} \mid cX_{13}X_{6,1} \mid bX_{13}X_{6,1} \mid \\ &\quad cX_{6,1} \mid bX_{6,1} \mid X_{11}X_7X_{6,1} \mid X_{11}X_{14}X_{6,1} \mid X_9X_{14}X_{6,1} \mid dX_{6,1} \mid X_{10}X_{14}X_{6,1} \\ X_{6,1} &\rightarrow X_{12}X_9 \mid X_{12}X_{13} \mid X_{12} \mid X_{12}X_9X_{6,1} \mid X_{12}X_{13}X_{6,1} \mid X_{12}X_{6,1} \\ X_7 &\rightarrow X_9X_{12} \\ X_8 &\rightarrow c \mid cX_{12} \mid bX_{12} \mid X_{11}X_7X_{12} \mid X_{11}X_{14}X_{12} \mid X_9X_{14}X_{12} \mid dX_{12} \mid X_{10}X_{14}X_{12} \mid b \mid cX_{8,1} \mid cX_{12}X_{8,1} \mid bX_{12}X_{8,1} \mid \\ &\quad X_{11}X_7X_{12}X_{8,1} \mid X_{11}X_{14}X_{12}X_{8,1} \mid X_9X_{14}X_{12}X_{8,1} \mid dX_{12}X_{8,1} \mid X_{10}X_{14}X_{12}X_{8,1} \mid bX_{8,1} \\ X_{8,1} &\rightarrow X_{13} \mid X_9X_{12} \mid X_{13}X_{12} \mid X_{13}X_{8,1} \mid X_9X_{12}X_{8,1} \mid X_{13}X_{12}X_{8,1} \\ X_9 &\rightarrow X_{11}X_7 \mid X_{11}X_{14} \mid d \mid X_{11}X_7X_{9,1} \mid X_{11}X_{14}X_{9,1} \mid dX_{9,1} \\ X_{9,1} &\rightarrow X_{14} \mid X_{14}X_{9,1} \\ X_{10} &\rightarrow X_{11}X_2 \mid c \\ X_{11} &\rightarrow a \\ X_{12} &\rightarrow b \\ X_{13} &\rightarrow c \\ X_{14} &\rightarrow d \end{aligned}$$

3. En nuestro tercer paso, eliminaremos todas las producciones de la forma  $X_i \rightarrow X_j\alpha$ , en donde  $i$  es mayor que  $j$ . Para hacer esto, reemplazaremos a  $X_j$  por todas sus producciones. Si al reemplazar aparecen otras producciones de esta forma, repetimos este paso, hasta que no queden más producciones en donde la primera variable de la producción tenga un índice menor al de la variable de la cabeza de la producción. En nuestro caso, no hay ninguna producción de este tipo, pero vamos a mostrar unas producciones de ejemplo para ilustrar el proceso:

$$\begin{aligned} X_9 &\rightarrow X_{11}X_7 \mid X_{11}X_{14} \mid d \mid X_{11}X_7X_{9,1} \mid X_{11}X_{14}X_{9,1} \mid dX_{9,1} \\ X_{10} &\rightarrow X_{11}X_2 \mid c \mid X_9X_{11}X_6 \end{aligned}$$

En este ejemplo, tenemos la producción  $X_{10} \rightarrow X_9X_{11}X_6$  en la que la variable de la izquierda de la producción ( $X_9$ ) tiene un índice menor al de la variable de la cabeza de la producción ( $X_{10}$ ). Reemplazando  $X_9$  por sus correspondientes producciones nos queda:



$$X_{10} \rightarrow X_{11}X_2 \mid c \mid X_{11}X_7X_{11}X_6 \mid X_{11}X_{14}X_{11}X_6 \mid dX_{11}X_6 \mid X_{11}X_7X_{9,1}X_{11}X_6 \mid X_{11}X_{14}X_{9,1}X_{11}X_6 \mid dX_{9,1}X_{11}X_6$$

4. En el último paso, reemplazamos todas las producciones en las que haya una variable al principio de la producción (no importa el índice que tenga) por las producciones de dicha variable que empiecen con algún símbolo terminal. Este proceso se realiza empezando por las últimas variables de la gramática y luego se continúa el proceso de reemplazo hacia arriba hasta llegar a las producciones de la variable de inicio. En nuestra gramática, partiríamos con las producciones de la variable  $X_{10}$  (ya que  $X_{11}$ ,  $X_{12}$ ,  $X_{13}$  y  $X_{14}$  tienen producciones en las que sólo hay involucrados símbolos terminales). La cosa nos quedaría así:

$$X_{10} \rightarrow X_{11}X_2 \mid c \Rightarrow X_{10} \rightarrow aX_2 \mid c$$

Luego continuaríamos con  $X_{9,1}$ :

$$X_{3,1} \rightarrow X_{14} \mid X_{14}X_{9,1} \Rightarrow X_{9,1} \rightarrow d \mid dX_{9,1}$$

El proceso no presenta mayor dificultad, hasta que llegamos a  $X_{8,1}$ , en donde aparecen las producciones  $X_{8,1} \rightarrow X_9X_{12}$  y  $X_{8,1} \rightarrow X_9X_{12}X_{8,1}$ , que no comienzan con  $X_{11}$ ,  $X_{12}$ ,  $X_{13}$  o  $X_{14}$ . Pero no importa, ya que todas las producciones de  $X_9$  para estas alturas del proceso han sido modificadas para que empiecen con un símbolo terminal, así que reemplazamos  $X_9$  directamente con sus producciones:

Esto:

$$X_{3,1} \rightarrow X_{13} \mid X_9X_{12} \mid X_{13}X_{12} \mid X_{13}X_{8,1} \mid X_9X_{12}X_{8,1} \mid X_{13}X_{12}X_{8,1}$$

Queda como esto:

$$X_{3,1} \rightarrow c \mid aX_7X_{12} \mid aX_{14}X_{12} \mid dX_{12} \mid aX_7X_{3,1}X_{12} \mid aX_{14}X_{9,1}X_{12} \mid dX_{9,1}X_{12} \mid cX_{12} \mid cX_{8,1} \mid aX_7X_{12}X_{8,1} \mid aX_{14}X_{12}X_{8,1} \mid dX_{12}X_{8,1} \mid aX_7X_{9,1}X_{12}X_{8,1} \mid aX_{14}X_{9,1}X_{12}X_{8,1} \mid dX_{9,1}X_{12}X_{8,1} \mid cX_{12}X_{8,1}$$

El proceso continúa así sin mayores complicaciones, hasta que llegamos a la siguiente gramática en forma normal de Greibach:

$$\begin{aligned} X_1 &\rightarrow aX_2 \mid c \mid cX_9X_3 \mid bX_9X_3 \mid cX_{13}X_3 \mid bX_{13}X_3 \mid cX_3 \mid bX_3 \mid aX_7X_3 \mid aX_{14}X_3 \mid aX_7X_{14}X_3 \mid aX_{14}X_{14}X_3 \mid dX_{14}X_3 \mid \\ &aX_7X_{9,1}X_{14}X_3 \mid aX_{14}X_{9,1}X_{14}X_3 \mid dX_{9,1}X_{14}X_3 \mid dX_3 \mid aX_2X_{14}X_3 \mid cX_{14}X_3 \mid cX_9X_6X_3 \mid bX_9X_{6,1}X_3 \mid cX_3X_{6,1}X_3 \mid \\ &bX_{13}X_{6,1}X_3 \mid cX_{6,1}X_3 \mid bX_{6,1}X_3 \mid aX_7X_{6,1}X_3 \mid aX_{14}X_{6,1}X_3 \mid aX_7X_{14}X_{6,1}X_3 \mid aX_{14}X_{14}X_{6,1}X_3 \mid dX_{14}X_{6,1}X_3 \mid \\ &aX_7X_{9,1}X_{14}X_{6,1}X_3 \mid aX_{14}X_{9,1}X_{14}X_{6,1}X_3 \mid dX_{9,1}X_{14}X_{6,1}X_3 \mid dX_{6,1}X_3 \mid aX_2X_{14}X_{6,1}X_3 \mid cX_{14}X_{6,1}X_{14}X_{6,1}X_3 \mid cX_9X_{11} \mid \\ &bX_9X_{11} \mid cX_{13}X_{11} \mid bX_{13}X_{11} \mid cX_{11} \mid bX_{11} \mid aX_7X_{11} \mid aX_{14}X_{11} \mid aX_7X_{14}X_{11} \mid aX_{14}X_{14}X_{11} \mid dX_{14}X_{11} \mid aX_7X_{9,1}X_{14}X_{11} \mid \\ &aX_{14}X_{9,1}X_{14}X_{11} \mid dX_{9,1}X_{14}X_{11} \mid dX_{11} \mid aX_2X_{14}X_{11} \mid cX_{14}X_{11} \mid cX_9X_{6,1}X_{11} \mid bX_9X_{6,1}X_{11} \mid cX_{13}X_{6,1}X_{11} \mid \\ &bX_{13}X_{6,1}X_{11} \mid cX_{6,1}X_{11} \mid bX_{6,1}X_{11} \mid aX_7X_{6,1}X_{11} \mid aX_{14}X_{6,1}X_{11} \mid aX_7X_{14}X_{6,1}X_{11} \mid aX_{14}X_{14}X_{6,1}X_{11} \mid dX_{14}X_{6,1}X_{11} \mid \\ &aX_7X_{9,1}X_{14}X_{6,1}X_{11} \mid aX_{14}X_{9,1}X_{14}X_{6,1}X_{11} \mid dX_{9,1}X_{14}X_{6,1}X_{11} \mid dX_{6,1}X_{11} \mid aX_2X_{14}X_{6,1}X_{11} \mid cX_{14}X_{6,1}X_{14}X_{6,1}X_{11} \mid cX_{11} \mid \\ &cX_{12}X_{11} \mid bX_{12}X_{11} \mid aX_7X_{12}X_{11} \mid aX_{14}X_{12}X_{11} \mid aX_7X_{14}X_{12}X_{11} \mid aX_{14}X_{14}X_{12}X_{11} \mid dX_{14}X_{12}X_{11} \mid aX_7X_{9,1}X_{14}X_{12}X_{11} \mid \\ &aX_{14}X_{9,1}X_{14}X_{12}X_{11} \mid dX_{9,1}X_{14}X_{12}X_{11} \mid dX_{12}X_{11} \mid aX_2X_{14}X_{12}X_{11} \mid cX_{14}X_{12}X_{11} \mid bX_{11} \mid cX_{8,1}X_{11} \mid cX_{12}X_{8,1}X_{11} \mid \\ &bX_{12}X_{8,1}X_{11} \mid aX_7X_{12}X_{8,1}X_{11} \mid aX_{14}X_{12}X_{8,1}X_{11} \mid aX_{14}X_{12}X_{8,1}X_{11} \mid dX_{12}X_{8,1}X_{11} \mid aX_2X_{14}X_{12}X_{8,1}X_{11} \mid cX_{14}X_{12}X_{8,1}X_{11} \mid \\ &bX_{8,1}X_{11} \mid a \mid aX_4 \mid aX_5 \\ X_2 &\rightarrow aX_2X_{12} \mid cX_{12} \\ X_3 &\rightarrow cX_{11} \mid cX_{12}X_{11} \mid bX_{12}X_{11} \mid aX_7X_{12}X_{11} \mid aX_{14}X_{12}X_{11} \mid aX_7X_{14}X_{12}X_{11} \mid aX_{14}X_{14}X_{12}X_{11} \mid dX_{14}X_{12}X_{11} \mid \\ &aX_7X_{9,1}X_{14}X_{12}X_{11} \mid aX_{14}X_{9,1}X_{14}X_{12}X_{11} \mid dX_{9,1}X_{14}X_{12}X_{11} \mid dX_{12}X_{11} \mid aX_2X_{14}X_{12}X_{11} \mid cX_{14}X_{12}X_{11} \mid bX_{11} \mid cX_{8,1}X_{11} \mid \\ &cX_{12}X_{8,1}X_{11} \mid bX_{12}X_{8,1}X_{11} \mid aX_7X_{12}X_{8,1}X_{11} \mid aX_{14}X_{12}X_{8,1}X_{11} \mid aX_{14}X_{12}X_{8,1}X_{11} \mid dX_{12}X_{8,1}X_{11} \mid aX_2X_{14}X_{12}X_{8,1}X_{11} \mid \\ &cX_{14}X_{12}X_{8,1}X_{11} \mid bX_{8,1}X_{11} \\ X_4 &\rightarrow aX_7X_5 \mid aX_{14}X_5 \mid dX_5 \mid aX_7X_{9,1}X_5 \mid aX_{14}X_{9,1}X_5 \mid dX_{9,1}X_5 \\ X_5 &\rightarrow cX_{14} \\ X_6 &\rightarrow cX_9 \mid bX_9 \mid cX_{13} \mid bX_{13} \mid c \mid b \mid aX_7 \mid aX_{14} \mid aX_7X_{14} \mid aX_{14}X_{14} \mid dX_{14} \mid aX_7X_{9,1}X_{14} \mid aX_{14}X_{9,1}X_{14} \mid dX_{9,1}X_{14} \mid d \mid \\ &aX_2X_{14} \mid cX_{14} \mid cX_9X_{6,1} \mid bX_9X_{6,1} \mid cX_{13}X_{6,1} \mid bX_{13}X_{6,1} \mid cX_{6,1} \mid bX_{6,1} \mid aX_7X_{6,1} \mid aX_{14}X_{6,1} \mid aX_7X_{14}X_{6,1} \mid \\ &aX_{14}X_{14}X_{6,1} \mid dX_{14}X_{6,1} \mid aX_7X_{9,1}X_{14}X_{6,1} \mid aX_{14}X_{9,1}X_{14}X_{6,1} \mid dX_{9,1}X_{14}X_{6,1} \mid dX_{6,1} \mid aX_2X_{14}X_{6,1} \mid cX_{14}X_{6,1}X_{14}X_{6,1} \\ X_{6,1} &\rightarrow bX_9 \mid bX_{13} \mid b \mid bX_9X_{6,1} \mid bX_{13}X_{6,1} \mid bX_{6,1} \\ X_7 &\rightarrow aX_7X_{12} \mid aX_{14}X_{12} \mid dX_{12} \mid aX_7X_{9,1}X_{12} \mid aX_{14}X_{9,1}X_{12} \mid dX_{9,1}X_{12} \\ X_8 &\rightarrow c \mid cX_{12} \mid bX_{12} \mid aX_7X_{12} \mid aX_{14}X_{12} \mid aX_7X_{14}X_{12} \mid aX_{14}X_{14}X_{12} \mid dX_{14}X_{12} \mid aX_7X_{9,1}X_{14}X_{12} \mid aX_{14}X_{9,1}X_{14}X_{12} \mid \\ &dX_{9,1}X_{14}X_{12} \mid dX_{12} \mid aX_2X_{14}X_{12} \mid cX_{14}X_{12} \mid b \mid cX_{8,1} \mid cX_{12}X_{8,1} \mid bX_{12}X_{8,1} \mid aX_7X_{12}X_{8,1} \mid aX_{14}X_{12}X_{8,1} \mid \\ &aX_{14}X_{12}X_{0,1} \mid dX_{12}X_{0,1} \mid aX_2X_{14}X_{0,1} \mid cX_{14}X_{12}X_{0,1} \mid bX_{0,1} \\ X_{8,1} &\rightarrow c \mid aX_7X_{12} \mid aX_{14}X_{12} \mid dX_{12} \mid aX_7X_{9,1}X_{12} \mid aX_{14}X_{9,1}X_{12} \mid dX_{9,1}X_{12} \mid cX_{12} \mid cX_{8,1} \mid aX_7X_{12}X_{8,1} \mid aX_{14}X_{12}X_{8,1} \mid \\ &dX_{12}X_{8,1} \mid aX_7X_{9,1}X_{12}X_{8,1} \mid aX_{14}X_{9,1}X_{12}X_{8,1} \mid dX_{9,1}X_{12}X_{8,1} \mid cX_{12}X_{8,1} \\ X_9 &\rightarrow aX_7 \mid aX_{14} \mid d \mid aX_7X_{9,1} \mid aX_{14}X_{9,1} \mid dX_{9,1} \\ X_{9,1} &\rightarrow d \mid dX_{9,1} \\ X_{10} &\rightarrow aX_2 \mid c \\ X_{11} &\rightarrow a \\ X_{12} &\rightarrow b \\ X_{13} &\rightarrow c \\ X_{14} &\rightarrow d \end{aligned}$$