Teoría de Autómatas y Lenguajes Formales

Boletín de Autoevaluación 5: **¿Cómo se simplifica una Gramática de Contexto Libre?.**

1. Objetivos.

El objetivo de este boletín es ilustrar cómo proceder para simplificar gramáticas de contexto libre 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.

Una gramática nos proporciona un conjunto de *producciones* que nos permitirán obtener cadenas de un determinado lenguaje. En este juego hay pocas reglas: una producción nos indica cómo substituir una serie de símbolos *auxiliares* por otra cadena de símbolos; se empieza siempre por un símbolo auxiliar especial, el símbolo S, *start*, y el objetivo es formar cadenas formadas únicamente por símbolos *terminales* esto es, símbolos del alfabeto sobre el que construimos el lenguaje. Al substituir no hay cota sobre el número de veces que se aplica una producción y tampoco hay reglas que tengan una "preferencia" sobre otras.

Al diseñar la gramática nuestra principal preocupación es asegurar que se producen todas las cadenas. De hecho, puede ocurrir que en el proceso de diseño se introduzcan producciones que no facilitan la obtención de cadenas o que introducen pasos innecesarios en el proceso de derivación.

El objetivo al simplificar una gramática de contexto libre es obtener una gramática equivalente, pero en la que se asegura que cada derivación es útil, y que no obliga a aplicar ningún paso innecesario.

3. Pasos para simplificar una Gramática de Contexto Libre.

Cualquier lenguaje de contexto libre, *L*, puede ser generado por medio de una GCL, *G*, que cumpla las siguientes condiciones:

1. Cada símbolo (terminal o auxiliar) de G se emplea en la derivación de alguna cadena de *L*.
2. Si $\lambda \notin L$ entonces en el conjunto de producciones de G no existen *producciones vacías*, es decir, producciones de la forma $A \to \lambda$.
3. En el conjunto de producciones de G no existen *producciones unitarias*, es decir, producciones de la forma $A \to B$ donde $A, B \in \Sigma_A$.

Si se obtiene una gramática que cumpla estas tres condiciones se puede asegurar que en cada derivación que se realiza se introduce información relevante. ¿Cómo se puede asegurar que son ciertas en nuestra gramática cada una de estas tres condiciones?

1. Podemos asegurar que *"cada símbolo (terminal o auxiliar) de G se emplea en la derivación de alguna cadena de* L*"* si **eliminamos los símbolos inútiles**.
2. Podemos asegurar que *"si $\lambda \notin L$ entonces en el conjunto de producciones de G no existen producciones de la forma $A \rightarrow \lambda$"* si **eliminamos las producciones vacías**.
3. Podemos asegurar que *"en el conjunto de producciones de G no existen* producciones unitarias*"* si **eliminamos las producciones unitarias**.

Estos son los tres pasos que debemos seguir para asegurar que la gramática está simplificada. A continuación se desarrolla cada uno de ellos.

3.1. Procedimiento.

Para ilustrar el procedimiento se simplificará la gramática $G = \{\Sigma_A, \Sigma_T, S, P\}$, en la que $\Sigma_A = \{S, A, B, C, D, E, F, G\}$, $\Sigma_T = \{a, b, c\}$ y el conjunto de producciones, P, es el siguiente:

$$
\begin{array}{l}
S \rightarrow Aab \mid B \mid CSa \mid b \\
A \rightarrow aA \mid Cb \mid a \mid aBAE \\
B \rightarrow bB \mid aBC \mid F \mid \lambda \\
C \rightarrow CG \mid DC \\
D \rightarrow aCb \mid a \\
E \rightarrow aaE \mid bB \\
F \rightarrow aF \mid ab \\
G \rightarrow F
\end{array}
$$

3.1.1. Eliminación de símbolos inútiles.

Hay que dividirlo en dos pasos y, además, el orden en que se realizan es significativo. Primero hay que eliminar los **símbolos no derivables** y, después, los **símbolos no alcanzables**.

Eliminar los símbolos no derivables. Un símbolo auxiliar que nunca se podrá substituir por un terminal (o una cadena de terminales) es un símbolo no derivable. Y se puede eliminar de la gramática: es inútil puesto que *si lo usamos nunca derivaremos una cadena formada exclusivamente por terminales*.

Una forma de actuar es calcular qué símbolos sí son derivables y desechar los que no lo sean, ya que la teoría nos da un procedimiento para determinar el conjunto de símbolos derivables:

Paso Base: $\forall\, A \in \Sigma_A,\ \forall\, w \in \Sigma_T^*$ tal que $A \rightarrow w \in P$, entonces se sabe que A es derivable.

Paso Recursivo: Si $(A \rightarrow \alpha) \in P$ y si todos los símbolos auxiliares de α son derivables, entonces el símbolo A también es derivable.

Siguiendo el algoritmo anterior, tendríamos que comenzar el proceso incluyendo en el conjunto de símbolos derivables, todos los auxiliares que tengan una producción que pertenezca a Σ_T^*, es decir, que se deriven en λ, en un terminal o en una subcadena de terminales.

Una vez determinado este conjunto inicial, el siguiente paso consiste en añadir como nuevos elementos los símbolos auxiliares que tengan al menos una producción en la que todos los símbolos son derivables. Al finalizar este paso, se tiene un nuevo conjunto de derivables; se vuelve a aplicar este paso, de nuevo se da un repaso a todas las producciones a ver si con este nuevo conjunto de derivables se puede añadir un nuevo auxiliar porque entre sus producciones hay alguna con todos sus símbolos derivables. El paso se repite sucesivamente, trabajando cada vez con el nuevo conjunto obtenido, hasta que en algún paso no se añada ningún nuevo elemento. Fíjate en el ejemplo.

Ejemplo:

Al partir del siguiente conjunto de producciones,

$$\begin{array}{l} S \to Aab \mid B \mid CSa \mid b \\ A \to aA \mid Cb \mid a \mid aBAE \\ B \to bB \mid aBC \mid F \mid \lambda \\ C \to CG \mid DC \\ D \to aCb \mid a \\ E \to aaE \mid bB \\ F \to aF \mid ab \\ G \to F \end{array}$$

se puede iniciar el conjunto de derivables, $\mathcal{D}$, con los siguientes símbolos:

$$\mathcal{D} = \{S, A, B, D, F\}$$

ya que al menos una producción de cada uno de esos símbolos auxiliares está formada por una cadena de Σ_T^*: $S \to b$, $A \to a$, $B \to \lambda$, $D \to a$ y $F \to ab$.

Aplicamos un primer paso recursivo, añadiendo los símbolos que contienen al menos una producción formada exclusivamente por derivables y que aún no pertenecen al conjunto $\mathcal{D}$; esto es, E (porque se tiene la producción $E \to bB$) y G (por la producción $G \to F$). Con esto, se tiene

$$\mathcal{D} = \{S, A, B, D, F, E, G\}$$

Se intenta aplicar otra vez el paso recursivo, pero no es posible añadir más símbolos: sólo falta C y no puede ser incluido, ya que ninguna de sus producciones está formada *exclusivamente* por derivables. Por lo tanto, *C no es derivable* y debe eliminarse *completamente* de la gramática. Se obtiene, por lo tanto, la gramática equivalente $G' = \{\Sigma_A', \Sigma_T, S, P'\}$, en la que $\Sigma_A' = \{S, A, B, D, E, F, G\}$, $\Sigma_T = \{a, b, c\}$ y el conjunto de producciones, P', es el siguiente:

$$\begin{array}{l} S \to Aab \mid B \mid b \\ A \to aA \mid a \mid aBAE \\ B \to bB \mid F \mid \lambda \\ D \to a \\ E \to aaE \mid bB \\ F \to aF \mid ab \\ G \to F \end{array}$$

que no tiene símbolos no derivables.

Eliminar los símbolos no alcanzables. Un símbolo auxiliar que nunca se podrá alcanzar desde el símbolo inicial de la gramática es un símbolo no alcanzable, esto es, un símbolo que nunca aparecerá en la derivación de una cadena porque no hay una secuencia de substituciones en las producciones que hagan que aparezca. Y, por supuesto, si nunca se va a usar, se puede eliminar de la gramática. El mismo razonamiento se puede aplicar a los símbolos terminales que aparezcan en la definición de la gramática pero que luego no aparecen en *ninguna* producción.

Como en el caso anterior, la forma más simple de desechar estos símbolos inútiles es determinando cuáles sí son alcanzables y eliminando lo que no lo sean. También en este caso la teoría brinda un procedimiento recursivo para calcular el conjunto de símbolos alcanzables:

Paso Base: $\Sigma'_A = \{S\}, \Sigma'_T = \emptyset$,

Paso Recursivo: si $A \in \Sigma'_A$ y $(A \to \alpha) \in P$, todos los símbolos auxiliares de α se añaden a Σ'_A y todos sus símbolos terminales se añaden a Σ'_T.

De acuerdo a esto, se inicia el conjunto de símbolos auxiliares alcanzables con S (y el conjunto de símbolos terminales útiles al conjunto vacío). En un primer paso, se incluyen en el conjunto de alcanzables todos los auxiliares que aparecen en producciones de S (y, en el conjunto de terminales útiles, los terminales). Con este nuevo conjunto, se realiza un segundo paso, similar, pero ahora incluyendo todos los auxiliares que aparecen en producciones de los alcanzables (y, en el conjunto de terminales útiles, los terminales). Cuando ya no se pueda añadir ningún nuevo elemento, finaliza el proceso y se eliminan los auxiliares que no estén incluidos en el conjunto de alcanzables (y, por lo que respecta a los terminales útiles, habría que redefinir la gramática indicando cuál es el conjunto de terminales "de verdad").

Ejemplo:

Si se parte de la gramática sin símbolos no derivables G', que se acaba de obtener, el valor inicial de los conjuntos de auxiliares alcanzables y de los terminales útiles se inician como:

$$\mathcal{A} = \{S\},\ \mathcal{T} = \emptyset$$

tal y como indica el método teórico.

En la primera aplicación del paso recursivo, se añaden a $\mathcal{A}$ los auxiliares A y B, y a $\mathcal{T}$ los terminales a y b; por lo tanto,

$$\mathcal{A} = \{S, A, B\},\ \mathcal{T} = \{a, b\}$$

En la segunda aplicación de este paso, hay que añadir los símbolos "que se alcanzan" desde A y desde B, que son E y F. No hay terminales nuevos que añadir a $\mathcal{T}$. Por lo tanto,

$$\mathcal{A} = \{S, A, B, E, F\},\ \mathcal{T} = \{a, b\}$$

En al tercera aplicación del paso recursivo, hay que añadir los auxiliares y terminales que intervienen en las producciones de E y F. Vemos que no se añaden símbolos nuevos, de lo que se deduce que D y G *no son alcanzables* y que el terminal c *nunca* se utiliza. Por lo tanto, la gramática equivalente a la original *sin símbolos inútiles* será:

$G'' = \{\Sigma''_A, \Sigma'_T, S, P''\}$, en la que $\Sigma''_A = \{S, A, B, E, F\}$, $\Sigma'_T = \{a, b\}$ y el conjunto de producciones, P'', es el siguiente:

$$\begin{array}{l} S \rightarrow Aab \mid B \mid b \\ A \rightarrow aA \mid a \mid aBAE \\ B \rightarrow bB \mid F \mid \lambda \\ E \rightarrow aaE \mid bB \\ F \rightarrow aF \mid ab \end{array}$$

3.1.2. Eliminación de las producciones vacías.

Cuando una gramática contiene producciones vacías, por ejemplo $A \rightarrow \lambda$, y en el proceso de generación de una cadena aparece el símbolo A en un forma sentencial, esto quiere decir que, más pronto o más tarde, el auxiliar A va a ser substituido por λ. Esto no supone un "avance" en el proceso de generar la cadena ya que substituir A por λ es lo mismo que eliminar A y, entonces ¿qué utilidad ha reportado el uso del auxiliar A en el proceso? Ninguna, y de hecho lo hemos "eliminado" al substituirlo por λ.

Por lo tanto, eliminar las producciones vacías es una forma de ahorrar trabajo inútil en el proceso de generación de una cadena. Para conseguirlo, el primer paso es determinar qué símbolos auxiliares son anulables, es decir, que símbolos pueden producir la cadena vacía en uno o más pasos. También en este caso se dispone de un algoritmo recursivo:

Paso Base: $\forall\ X \in \Sigma_A \mid (X \rightarrow \lambda) \in P$, entonces se sabe que X es anulable

Paso Recursivo: Si $(X \rightarrow \alpha) \in P$ y si todos los símbolos auxiliares de α son anulables, entonces X también es anulable.

El proceso finaliza cuando no se puedan añadir nuevos valores al conjunto de símbolos anulables.

Ejemplo:

Si seguimos con el ejemplo anterior, y aplicamos el procedimiento teórico, el conjunto de símbolos anulables, $\mathcal{N}$, tendrá como valor inicial

$$\mathcal{N} = \{B\}$$

ya que B es el único auxiliar que tiene una producción vacía. Al aplicar por primera vez el paso recursivo se añade S, ya que tiene una producción formada exclusivamente por anulables, $S \rightarrow B$,

$$\mathcal{N} = \{S, B\}$$

Si se aplica una segunda vez el paso recursivo no se añade ningún nuevo símbolo, por lo que el proceso ya finaliza.

Una vez que ya está calculado el conjunto de símbolos anulables, hay que reescribir las producciones de la gramática. Cada producción se reescribe eliminando ninguna, una o más apariciones de cada elemento del conjunto de símbolos anulables que aparezca en el consecuente de la producción. Expresado de otra forma, cada producción con símbolos anulables se reescribe varias veces: una tal y como está expresada originalmente y todas las necesarias para indicar todas las combinaciones posibles de anulables que "desaparecen". Con el ejemplo se ve más claro.

Ejemplo: En este punto del proceso, las producciones de la gramática son:

$$\begin{array}{l} S \rightarrow Aab \mid B \mid b \\ A \rightarrow aA \mid a \mid aBAE \\ B \rightarrow bB \mid F \mid \lambda \\ E \rightarrow aaE \mid bB \\ F \rightarrow aF \mid ab \end{array}$$

y, tal y como hemos visto, $\mathcal{N} = \{S, B\}$. ¿Cómo reescribir las producciones? Comencemos con las de los símbolos no anulables, por ejemplo con las de A,

$$A \rightarrow aA \mid a \mid aBAE$$

en $A \rightarrow aA$ no intervienen anulables y en $A \rightarrow a$, tampoco. En $A \rightarrow aBAE$, evidentemente, sí. Esa producción debe reescribirse, una vez tal cual está y otra teniendo en cuenta que B puede ser λ:

$$A \rightarrow aA \mid a \mid aBAE \mid aAE$$

Pasemos a estudiar las producciones de E:

$$E \rightarrow aaE \mid bB$$

Hay que reescribir la producción $E \rightarrow bB$ que es la única que contiene un anulable. Se escribe tal cual y teniendo en cuenta que B puede ser λ,

$$E \rightarrow aaE \mid bB \mid b$$

Con las producciones de F no es necesario hacer una reescritura porque no contienen símbolos anulables.

Pasemos, a continuación, a ver qué se hace con las producciones de los símbolos anulables, por ejemplo B:

$$B \rightarrow bB \mid F \mid \lambda$$

En principio, con la producción $B \rightarrow bB$ se procede como con las demás: se escribe tal cual y teniendo en cuenta que B puede ser λ. Y este sería el resultado:

$$B \rightarrow bB \mid b \mid F$$

¿Qué ocurre? Además de reescribir la producción con símbolos anulables se ha eliminado la producción vacía; esto es así, *porque ya se ha tenido en cuenta al escribir las producciones que B podría ser* λ. Por lo tanto, ya no es necesaria esa producción.

Finalmente, se reescriben las producciones de S,

$$S \rightarrow Aab \mid B \mid b$$

En $S \rightarrow Aab$ no intervienen anulables y en $S \rightarrow b$, tampoco. En $S \rightarrow B$, evidentemente, sí. Esa producción debe reescribirse, una vez tal cual está y otra teniendo en cuenta que B puede ser λ:

$$S \rightarrow Aab \mid B \mid \lambda \mid b$$

¿Curioso, eh? Ha aparecido una producción vacía entre las producciones de S (lo que no debe sorprendernos porque S es anulable, no lo olvidemos). ¿Por qué no se ha eliminado esa producción tal y como se ha hecho con la producción vacía de B? Pues porque no se puede. Si el símbolo inicial se deriva en λ, eso quiere decir que $\lambda \in L(G)$ y si se elimina esa producción, la gramática obtenida no puede ser equivalente (los lenguajes son distintos, uno contiene a λ y el otro, no). Pero, por ahora, vamos a ignorar este hecho. Nos apuntamos que $S \rightarrow \lambda \in P$ y se elimina esa producción vacía, quedando las producciones como sigue:

$$\begin{array}{l} /*\ No\ olvidar\ que\ S \rightarrow \lambda */ \\ S \rightarrow Aab \mid B \mid b \\ A \rightarrow aA \mid a \mid aBAE \mid aAE \\ B \rightarrow bB \mid b \mid F \\ E \rightarrow aaE \mid bB \mid b \\ F \rightarrow aF \mid ab \end{array}$$

Como al eliminar las producciones λ no se han "reintroducido" símbolos inútiles, se pasa a la última etapa.

3.1.3. Eliminación de las producciones unitarias.

Las producciones de la forma $A \rightarrow B$ dan lugar a trabajo innecesario, ya que cuando aparecen en una derivación lo único que introducen es un cambio de nombre del auxiliar. Por eso también se eliminan cuando se pretende obtener una gramática simplificada.

El nuevo conjunto de producciones se forma de la manera siguiente: Si $(A \rightarrow B) \in P$ y $A, B \in \Sigma_A$, entonces esta producción se elimina del nuevo conjunto de producciones y se introducen las siguientes nuevas producciones: $(A \rightarrow \gamma) \ \mid \gamma$ es una forma sentencial que se obtiene a partir de B, al aplicar producciones de P. Es decir, a las producciones de A añadimos las producciones de B.

Ejemplo:

Tal y como está en este momento la gramática, con el siguiente conjunto de producciones

$$\begin{array}{l} /*\ No\ olvidar\ que\ S \rightarrow \lambda */ \\ S \rightarrow Aab \mid B \mid b \\ A \rightarrow aA \mid a \mid aBAE \mid aAE \\ B \rightarrow bB \mid b \mid F \\ E \rightarrow aaE \mid bB \mid b \\ F \rightarrow aF \mid ab \end{array}$$

sólo hay dos producciones unitarias, $S \rightarrow B$ y $B \rightarrow F$. Se reescriben las producciones de S substituyendo B por sus producciones (y "arreglándolas": si lo hago tal cual la producción $S \rightarrow b$ aparece dos veces.. ¿con una llega, no?),

$$\begin{array}{l} /*\ No\ olvidar\ que\ S \rightarrow \lambda */ \\ S \rightarrow Aab \mid bB \mid b \mid F \\ A \rightarrow aA \mid a \mid aBAE \mid aAE \\ B \rightarrow bB \mid b \mid F \\ E \rightarrow aaE \mid bB \mid b \\ F \rightarrow aF \mid ab \end{array}$$

y, ahora toca substituir F en las producciones de B... y en las de S, ya que al substituir B, se ha introducido F,

$$\begin{array}{l} /*\ No\ olvidar\ que\ S \rightarrow \lambda */ \\ S \rightarrow Aab \mid bB \mid b \mid aF \mid ab \\ A \rightarrow aA \mid a \mid aBAE \mid aAE \\ B \rightarrow bB \mid b \mid aF \mid ab \\ E \rightarrow aaE \mid bB \mid b \\ F \rightarrow aF \mid ab \end{array}$$

No se han “reintroducido” símbolos inútiles. Para tener la gramática simplificada equivalente a la original sólo hace falta ver qué hacer cuando, como en este caso, el símbolo inicial es anulable y, por lo tanto, $\lambda \in L(G)$. Recordad que hay un lema que da la solución:

> *“Dada una GCL G puede construirse una GCL G′ equivalente a G tal que no tenga producciones λ excepto cuando $\lambda \in L(G)$ en cuyo caso $S' \rightarrow \lambda$ es la única producción en la que aparece λ y además S′ no aparece en el consecuente de ninguna otra regla de producción.”*

Pues aplicamos el lema; se introduce un nuevo símbolo inicial, S′, y será el único auxiliar con una producción vacía. Y, para asegurar la segunda condición, se añade otra producción *unitaria* (la única... nos ha salido un poco transgresor este S';-) . La gramática quedará finalmente de la siguiente forma:

$G^s = \{\Sigma_A^s, \Sigma_T^s, S', P^s\}$, en la que $\Sigma_A^s = \{S', S, A, B, E, F\}$, $\Sigma_T^s = \{a, b\}$ y el conjunto de producciones, P^s, es el siguiente:

$$\begin{array}{l} S' \rightarrow \lambda \mid S \\ S \rightarrow Aab \mid bB \mid b \mid aF \mid ab \\ A \rightarrow aA \mid a \mid aBAE \mid aAE \\ B \rightarrow bB \mid b \mid aF \mid ab \\ E \rightarrow aaE \mid bB \mid b \\ F \rightarrow aF \mid ab \end{array}$$

3.2. Otro Ejemplo.

El ejemplo desarrollado en el apartado anterior era “completito” en el sentido de que nos fijábamos en la gramática (que es lo que se debe hacer desde el punto de vista formal). Pero lo habitual en este asignatura es que tengamos un enunciado como el siguiente:

Simplificar la siguiente GCL:

$$
\begin{aligned}
&S \rightarrow aSb \mid aAb \\
&A \rightarrow BA \mid \lambda \mid aCB \\
&B \rightarrow ASb \mid abB \\
&C \rightarrow aDF \mid aDb \\
&D \rightarrow abC \mid aCB \mid aF \\
&F \rightarrow Fb \mid aCb
\end{aligned}
$$

en el que prestamos atención principalmente al conjunto de producciones. Ello obliga, por ejemplo, a que no solamos plantearnos cuestiones como, por ejemplo, si todos los terminales son o no útiles (sólo trabajamos con los útiles, los que sí aparecen en la producciones).

Por eso, este ejemplo nos limitaremos al desarrollo habitual en clase, dejando el del apartado anterior como ejemplo *completo* del proceso de simplificación. Se plantean, entonces, cada uno de los tres pasos:

1. **Eliminación de símbolos inútiles:** En dos pasos, sin olvidar que el orden es significativo,

   a) <u>Eliminación de símbolos no derivables:</u>
   El valor inicial del conjunto de símbolos derivables, $\mathcal{D}$, será $\mathcal{D} = \{A\}$ ya que A es el único símbolo auxiliar con una producción en una cadena de Σ_T^*, $A \rightarrow \lambda$.
   Al aplicar una vez el paso recursivo, puede añadirse el símbolo S, ya que entre sus producciones está la producción $S \rightarrow aAb$ en la que todos los símbolos son derivables. Por lo tanto, tras este paso,

   $$\mathcal{D} = \{A, S\}.$$

   Si se vuelve a aplicar el paso recursivo, se puede añadir el símbolo B, ya que entre sus producciones está $B \rightarrow ASb$ en la que todos los símbolos son derivables. Así,

   $$\mathcal{D} = \{A, S, B\}.$$

   Si se vuelven a examinar las producciones, se puede comprobar que ya no es posible añadir ningún nuevo elemento a $\mathcal{D}$. Es decir,

   $$\mathcal{D} = \{S, A, B\}$$

   y los símbolos C, D y F son no derivables y se eliminan completamente de la gramática, quedando las producciones de la siguiente forma:

   $$
   \begin{aligned}
   &S \rightarrow aSb \mid aAb \\
   &A \rightarrow BA \mid \lambda \\
   &B \rightarrow ASb \mid abB
   \end{aligned}
   $$

   b) <u>Eliminación de símbolos no alcanzables:</u>
   Si sólo nos centramos en los símbolos auxiliares (al partir de las producciones, asumimos que $\Sigma_T = \{a, b\}$ y que todos los terminales son útiles), el paso base indica el siguiente valor inicial para $\mathcal{A}$,

   $$\mathcal{A} = \{S\}$$

Al aplicar el paso recursivo, añadimos A, ya que $S \to aAb$,

$$\mathcal{A} = \{S, A\}$$

y si se vuelve a aplicar, se debe añadir B, ya que $A \to BA$,

$$\mathcal{A} = \{S, A, B\}$$

No hay más símbolos auxiliares. Por lo tanto, todos los símbolos auxiliares son alcanzables y no se elimina ninguno.

2. **Eliminación de producciones vacías:** Sólo hay una producción vacía, $A \to \lambda$. Por lo tanto, el valor incial de $\mathcal{N}$ es

$$\mathcal{N} = \{A\}$$

Al aplicar el paso recursivo, se ve que ningún otro símbolo tiene producciones formadas exclusivamente por símbolos anulables. Por lo tanto, no se añade ningún nuevo elemento a $\mathcal{N}$ y A resulta ser el único símbolo anulable. Por lo tanto, se va a eliminar su producción vacía teniendo en cuenta que cada producción en la que aparezca A se debe escribir tal cual y teniendo en cuenta que A puede ser λ:

$$\begin{array}{l} S \to aSb \mid aAb \mid ab \\ A \to BA \mid B \\ B \to ASb \mid Sb \mid abB \end{array}$$

No se han introducido símbolos inútiles en el proceso de eliminar las producciones vacías, por lo que se pasa a la siguiente etapa.

3. **Eliminación de producciones unitarias:** Sólo aparece una producción unitaria, $A \to B$; se substituye B por sus producciones

$$\begin{array}{l} S \to aSb \mid aAb \mid ab \\ A \to BA \mid ASb \mid Sb \mid abB \\ B \to ASb \mid Sb \mid abB \end{array}$$

y, como no hay más producciones unitarias, ni se han introducido símbolos inútiles en el proceso, se puede asegurar ya que la gramática está simplificada.

4. Autoevaluación.

1. Simplificar la siguiente GCL

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

<u>Solución:</u>

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

2. Simplificar la siguiente GCL

$$\begin{array}{l} S \to aAb \mid cHB \mid CH \\ A \to dBH \mid eeC \\ B \to ff \mid D \\ C \to gFB \mid ah \\ D \to i \\ E \to jF \\ F \to dcGGG \mid cF \\ G \to kF \\ H \to Hlm \end{array}$$

<u>Solución:</u>

$$\begin{array}{l} S \to aAb \\ A \to eeC \\ C \to ah \end{array}$$

3. Simplificar la siguiente GCL

$$\begin{array}{l} S \to aAB \mid bBA \mid ABb \mid SS \\ A \to aAb \mid CCA \mid BB \\ B \to \lambda \mid bC \\ C \to aCS \mid SCS \\ D \to ab \mid SABC \end{array}$$

<u>Solución:</u>

$$\begin{array}{l} S \to aA \mid a \mid bA \mid b \mid Ab \mid SS \\ A \to aAb \mid ab \end{array}$$

4. Simplificar la siguiente gramática:

$$\begin{array}{l} S \to CBa \mid D \\ A \to bbC \\ B \to Sc \mid ddd \\ C \to eA \mid f \mid C \\ D \to E \mid SABC \\ E \to gh \end{array}$$

<u>Solución:</u>

$$\begin{array}{l} S \to CBa \mid gh \mid SABC \\ A \to bbC \\ B \to Sc \mid ddd \\ C \to eA \mid f \end{array}$$