

Uma Gramática Livre de Contexto pode ser representada de acordo com algumas Formas Normais, que são modos de representar uma gramática visando a otimização na geração de palavras da linguagem. Duas destas formas podem ser destacadas, a Forma Normal de Chomsky (FNC) e a Forma Normal de Greibach (FNG), na qual a FNC visa a cada derivação gerar duas variáveis ou somente um terminal, enquanto que a FNG visa a cada derivação gerar pelo menos um terminal, seguido de somente variáveis. Ambas as formas normais são conteúdos a serem vistos na sequência deste texto.

### **Forma Normal de Chomsky**

Uma GLC é dita estar na Forma Normal de Chomsky (FNC) quando todas as suas regras de produção estão representadas no formato  $A \rightarrow BC$  ou  $A \rightarrow a$ , sendo que  $A, B$  e  $C$  são elementos de  $V$ , e  $a$  é um elemento de  $T$ . Assim, uma gramática na FNC é descrita de tal modo que todas as suas regras de produção, ou geram duas variáveis, ou geram um terminal, possibilitando assim uma árvore de derivação binária, na qual um nó ou tem dois filhos (duas variáveis), ou um único nó filho (terminal).

Dada uma gramática livre de contexto, pode-se representar como uma gramática na Forma Normal de Chomsky utilizando-se de um algoritmo que converta a gramática para FNC. Este algoritmo pode ser representado por três etapas: simplificação da gramática (eliminando regras de produções vazias, regras de produções unitárias e símbolos inúteis), transformação de símbolos terminais em variáveis e transformação do lado direito das regras de produção em tamanho igual a dois. Para a primeira etapa de simplificação da GLC, realizam-se os mesmos passos vistos na simplificação.

A segunda etapa tem como objetivo deixar o lado direito das regras de produção que não estão no formato, representadas somente com variáveis. Para isso são criadas novas variáveis para os terminais que aparecem nas regras que não estão no formato, e estas são substituídas por estas novas variáveis.

Por fim, a terceira etapa visa transformar as regras de produção cujo lado direito tem tamanho maior que três em tamanho exatamente dois, uma vez que o formato da FNC tem tamanho dois e com duas variáveis ( $A \rightarrow BC$ ). Isto é feito novamente criando novas variáveis auxiliares para reduzir o tamanho das regras de produção.

Um exemplo de transformação de uma GLC na FNC pode ser visualizado no exemplo para a gramática abaixo:

$G = \{V = \{S, A, B, C, D, E\}, T = \{a, b, c, d\}, P, S\}$ , onde:

$P = \{ S \rightarrow ABB \mid CB \mid BE$

$A \rightarrow aaA \mid \lambda$

$B \rightarrow bBb \mid A$

$$C \rightarrow ccC$$

$$D \rightarrow dD \mid d\}$$

Após a gramática acima passar pela etapa de simplificação (vazias, unitárias e inúteis), tem-se a seguinte gramática resultante:

$G = \{V = \{S, A, B\}, T = \{a, b\}, P, S\}$ , onde:

$$P = \{ S \rightarrow ABB \mid BB$$

$$A \rightarrow aaA \mid aa$$

$$B \rightarrow bBb \mid bb \mid aaA \mid aa\}$$

Na segunda etapa, as regras  $S \rightarrow ABB$  e  $S \rightarrow BB$  não precisam ser alteradas pois não existem terminais nestas regras.

Como a segunda etapa da FNC tem como objetivo transformar os terminais das regras que não estão no formato em variáveis, as seguintes regras de produção devem ser alteradas:  $A \rightarrow aaA$ ,  $A \rightarrow aa$ ,  $B \rightarrow bBb$ ,  $B \rightarrow bb$ ,  $B \rightarrow aaA$  e  $B \rightarrow aa$ . Para estas regras, podemos observar que existem dois terminais diferentes, a e b. Para estes terminais devem ser criados uma variável para cada, sendo que cada variável terá uma única regra de produção, gerando exclusivamente um terminal, assim podemos definir uma nova variável X, que gera o terminal a, e uma nova variável Y, que gera o terminal b, resultando na adição de duas regras  $X \rightarrow a$  e  $Y \rightarrow b$ .

Definidas estas novas regras de X e Y, substituímos os terminais a e b, por estas variáveis X e Y nas regras que devem ser alteradas nesta etapa, assim a regra de produção  $A \rightarrow aaA$  resulta em  $A \rightarrow XXA$ ,  $A \rightarrow aa$  resulta em  $A \rightarrow XX$ ,  $B \rightarrow bBb$  transformada fica  $B \rightarrow YBY$ ,  $B \rightarrow bb$  resulta em  $B \rightarrow YY$ , a regra  $B \rightarrow aaA$  transforma-se em  $B \rightarrow XXA$  e por fim a regra  $B \rightarrow aa$  substituída resulta em  $B \rightarrow XX$ , resultando nas seguintes regras de produção:

$$P = \{ S \rightarrow ABB \mid BB$$

$$A \rightarrow XXA \mid XX$$

$$B \rightarrow YBY \mid YY \mid XXA \mid XX$$

$$X \rightarrow a$$

$$Y \rightarrow b\}$$

Algumas regras de produção acima já estão no formato das regras na FNC e não precisam ser alteradas nesta terceira etapa, porém para  $S \rightarrow ABB$ ,  $A \rightarrow XXA$ ,  $B \rightarrow$

$YBY$  e  $B \rightarrow XXA$  é necessário criar novas variáveis para deixá-las com somente duas variáveis do lado direito das regras de produção.

Para a regra  $S \rightarrow ABB$ , se criarmos uma nova variável  $M \rightarrow BB$ , podemos substituir na regra  $S \rightarrow ABB$  por  $S \rightarrow AM$ , adicionando a regra  $M$  as regras de produções e deixando ambas no formato da FNC.

A regra de produção  $A \rightarrow XXA$ , podemos criar uma nova variável  $N$  para  $XX$ , resultando em  $N \rightarrow XX$ , e substituindo ficamos com  $A \rightarrow NA$ . Esta mesma substituição pode ser feita para a regra  $B \rightarrow XXA$ , resultando em  $B \rightarrow NA$ , e mantendo a regra  $N \rightarrow XX$ .

Por fim a regra  $B \rightarrow YBY$ , podemos criar uma nova variável  $Q \rightarrow YB$ , e substituir na regra original, resultando em  $B \rightarrow QY$  e adicionando a regra de  $Q$  as regras de produção.

As regras resultantes após a terceira etapa então podem ser vistas abaixo, bem como a gramática resultante na Forma Normal de Chomsky.

$G = \{V = \{S, A, B, X, Y, M, N, Q\}, T = \{a, b\}, P, S\}$ , onde:

$P = \{ S \rightarrow AM \mid BB$

$A \rightarrow NA \mid XX$

$B \rightarrow QY \mid YY \mid NA \mid XX$

$X \rightarrow a$

$Y \rightarrow b$

$M \rightarrow BB$

$N \rightarrow XX$

$Q \rightarrow YB\}$

Podemos observar que na terceira etapa criamos a variável  $N \rightarrow XX$ , e substituímos nas regras de  $A \rightarrow XXA$  e  $B \rightarrow XXA$ . Ainda nas regras de produção temos uma regra de produção  $B \rightarrow XX$ . Lembrando que o  $N$  não pode substituir o  $XX$  na regra de  $B \rightarrow XX$ , pois teríamos uma regra  $B \rightarrow N$ , que estaria fora do formato de uma gramática na FNC, assim, na terceira etapa somente alteramos as regras que não estão no formato, criando as variáveis auxiliares para as regras de produção cujo tamanho do lado direito seja maior que dois.

Outro ponto importante relativo a FNC é a escolha de quais variáveis irão compor a nova variável, por exemplo, para  $S \rightarrow ABB$  escolhemos as duas últimas variáveis,  $BB$  para criar a variável auxiliar  $M \rightarrow BB$ . Para a regra de produção  $B \rightarrow YBY$ , escolhemos

as duas primeiras variáveis, YB para criar a variável auxiliar Q, definindo assim a regra de produção  $Q \rightarrow YB$ . Isto enfatiza o fato de que a escolha de quais variáveis irão compor a nova regra de produção pode ser feita de qualquer forma, agrupando sempre duas variáveis em uma nova variável auxiliar, gerando então uma nova regra de produção já no formato da FNC.

Exemplo 2: para a seguinte GLC, coloque-a na Forma Normal de Chomsky, e apresente a gramática resultante.

$G = \{V = \{S, A, B, C, D\}, T = \{a, b, c, d\}, P, S\}$ , onde:

$P = \{ S \rightarrow aABb \mid BC \mid CD$

$A \rightarrow aA \mid B$

$B \rightarrow bbB \mid \lambda$

$C \rightarrow Cc \mid dDd$

$D \rightarrow dD \mid cC\}$

Começando pela simplificação da gramática, temos como primeira etapa a criação do conjunto  $V_\lambda = \{B, A\}$ , na qual considerando que estas variáveis podem ser vazias, adicionamos novas regras de produção, resultando nas seguintes regras:

$P = \{ S \rightarrow aABb \mid ABb \mid aAb \mid ab \mid BC \mid C \mid CD$

$A \rightarrow aA \mid a \mid B$

$B \rightarrow bbB \mid bb$

$C \rightarrow Cc \mid dDd$

$D \rightarrow dD \mid cC\}$

Na segunda etapa da simplificação, a eliminação das regras de produções unitárias, tem-se duas regras que precisam ser alteradas,  $S \rightarrow C$  e  $A \rightarrow B$ , que resultam nas seguintes regras:

$P = \{ S \rightarrow aABb \mid ABb \mid aAb \mid ab \mid BC \mid Cc \mid dDd \mid CD$

$A \rightarrow aA \mid a \mid bbB \mid bb$

$B \rightarrow bbB \mid bb$

$C \rightarrow Cc \mid dDd$

$D \rightarrow dD \mid cC\}$

Por fim, na terceira etapa da simplificação, eliminamos as variáveis inalcançáveis, improdutivas ou as variáveis que fiquem em loop infinito, sendo que o resultado após esta etapa é apresentado abaixo:

$$P = \{ S \rightarrow aABb \mid ABb \mid aAb \mid ab$$

$$A \rightarrow aA \mid a \mid bbB \mid bb$$

$$B \rightarrow bbB \mid bb\}$$

Continuando na transformação da gramática na Forma Normal de Chomsky, como segundo passo, criamos uma variável para cada terminal que aparecer em regras que não estão no formato, sendo assim podemos utilizar o  $X \rightarrow a$  e  $Y \rightarrow b$ , e substituímos nas regras acima resultando nas seguintes regras de produção:

$$P = \{ S \rightarrow XABY \mid ABY \mid XAB \mid XY$$

$$A \rightarrow XA \mid a \mid YYB \mid YY$$

$$B \rightarrow YYB \mid YY$$

$$X \rightarrow a$$

$$Y \rightarrow b\}$$

Finalmente, a terceira etapa da transformação de uma GLC na FNC deixa as regras de produções com tamanho exatamente dois, de acordo com o formato. A gramática transformada na FNC então pode ser visualizada abaixo

$G = \{V = \{S, A, B, X, Y, M, N, Q\}, T = \{a, b\}, P, S\}$ , onde:

$$P = \{ S \rightarrow MN \mid AN \mid MB \mid XY$$

$$A \rightarrow XA \mid a \mid QB \mid YY$$

$$B \rightarrow QB \mid YY$$

$$X \rightarrow a$$

$$Y \rightarrow b$$

$$M \rightarrow XA$$

$$N \rightarrow BY$$

$$Q \rightarrow YY\}$$

## **Bibliografia**

- VIEIRA, NEWTON JOSÉ. **Introdução aos Fundamentos da Computação**. São Paulo. Pioneira Thomson Learning. 2006;
- MENEZES, P. B.; DIVERIO, T. A.; **Teoria da Computação: Máquinas Universais e Computabilidade**; 3ª edição Bookman 2011
- MENEZES, P. B. ; **Linguagens Formais e Autômatos**. 6ª edição. Ed. Artmed. 2011
- SIPSER M. **Introdução à Teoria da Computação**. 2 ed. Cengage Learning.2007
- MORET, B. M. "**Theory of Computation**". Addison-Wesley, 1998.
- HOPCROFT, JOHN E.; ULLMAN, JEFFREY D.; MOTWANI, RAJEEV **Introdução à Teoria de Autômatos, Linguagens e Computação** Ed.Campus 2002
- SILVA, FLAVIO SOARES CORRÊA; MELO, ANA CRISTINA VIEIRA; **Modelos Clássicos de Computação** Ed. Thomson 1ª Edição 2006