Formas Normais

As Formas Normais estabelecem restrições rígidas na forma das Produções, sem reduzir o poder de geração das gramáticas livres do contexto. (exceturando-se a geração da palavra vazia).

São usadas no desenvolvimento de algoritmos (reconhecedores de linguagens, por exemplo) e na prova de teoremas.

As formas normais apresentadas são as seguintes. Suponha que A, B e C são variáveis, a é terminal e α é uma palavra de variáveis.

* Forma Normal de Chomsky, na qual as produções são da forma:

$$A \to BC \quad \text{ou} \quad A \to a$$

* Forma Normal de Greibach, na qual as produções são da forma:

$$A \to a\alpha$$

Forma Normal de Chomsky

Uma GLC $G = (V, T, P, S)$ é dita estar na Forma Normal de Chomsky se todas as suas produções são da forma:

$$A \to BC \quad \text{ou} \quad A \to a$$

Suponha que A, B e C são variáveis de V e a é terminal de T.

Portanto, a palavra vazia não pertence à linguagem gerada por uma gramática na Formal Normal de Chomsky. (FNC).

O algoritmo a seguir transforma uma GLC qualquer, cuja linguagem gerada não possua a palavra vazia, em uma gramatica na FNC. O algoritmo é dividido em três etapas, como segue:

ⓐ ETAPA 1: Simplificação da gramática. Simplifica a gramática, excluindo as produções vazias (como a linguagem não possui a palavra vazia, todas as produções da forma $A \to \varepsilon$ podem ser excluídas), produções da forma $A \to B$ (se o lado direito de alguma produção tiver somente um símbolo, este será terminal) e, opcionalmente, os símbolos inúteis;

ETAPA 2: Transformação do lado direito das produções de comprimento maior ou igual a dois. (3)

Garante que o lado direito das produções de comprimento maior ou igual a dois é composto exclusivamente por variáveis. A exclusão de um terminal a pode ser realizada, substituindo-se este por uma variável intermediária C_a e incluindo a produção $C_a \rightarrow a$.

ETAPA 3: Transformação do lado direito das produções de comprimento maior ou igual a três, em produções com exatamente duas variáveis.

Garante que o lado direito das produções de comprimento maior que um é composto exatamente por duas variáveis. Após a execução da etapa acima, o lado direito das produções da forma $A \rightarrow B_1 B_2 \ldots B_n$ $(n \geq 2)$ é composto exclusivamente por variáveis. Portanto, para concluir a transformação, é suficiente garantir que o lado direito é composto por exatamente duas variáveis. Isto é possível, gerando-se $B_1 B_2 \ldots B_n$ em diversas etapas, e usando-se variáveis intermediárias.

Algoritmo: Forma Normal de Chomsky

Seja $G = (V, T, P, S)$ uma gramática livre de contexto tal que $\epsilon \notin GERA(G)$ ou $\epsilon \notin L(G)$. O algoritmo para transformar na FNC é como segue:

ETAPA 1: Simplificação da gramática. As seguintes simplificações:

* produções vazias;
* produções que substituem variáveis;
* símbolos inúteis (opcional);

Devem ser realizadas usando os algoritmos de simplificação descritos anteriormente, resultando na gramática:

$$G_1 = (V_1, T_1, P_1, S)$$

ETAPA 2: Transformação do lado direito das produções de comprimento maior ou igual a dois.

A gramática resultante desta etapa é:

$$G_2 = (V_2, T_1, P_2, S)$$

na qual V_2 e P_2 são construídos conforme o algoritmo: [para cada variável a, suponha $C_a \notin V_2$];

$V_2 = V_1$; $P_2 = P_1$;
para toda $A \to X_1 X_2 X_n \in P_2$ tal que $n \geq 2$.
faça se
 para $r \in \{1, \ldots, n\}$, X_r é um símbolo terminal
 então (suponha $X_r = a$)
 $V_2 = V_2 \cup \{C_a\}$
 substitui a por C_a em $A \to X_1 X_2 \ldots X_n \in P_2$
 $P_2 = P_2 \cup \{C_a \to a\}$;

ETAPA 3: Transformação do lado direito das produções de comprimento maior ou igual a três em produções com exatamente duas variáveis. A gramática resultante desta etapa é:

$$G_3 = (V_3, T_1, P_3, S)$$

na qual V_3 e P_3 são construídos conforme o algoritmo: [A cada ciclo, suponha $D_1 \notin V_3, \ldots D_{n-2} \notin V_3$]

$V_3 = V_2$;

$P_3 = P_2$;

para toda $A \to B_1 B_2 \ldots B_n \in P_3$ tal que $n \geq 3$

faça $P_3 = P_3 - \{A \to B_1 B_2 \ldots B_n\}$;

$V_3 = V_3 \cup \{D_1, \ldots, D_{n-2}\}$;

$P_3 = P_3 \cup \{A \to B_1 D_1, D_1 \to B_2 D_2, \ldots, D_{n-3} \to B_{n-2} D_{n-2}, D_{n-2} \to B_{n-1} B_n\}$;

Exemplo: Considere a Gramática.

$G = (\overset{V_1}{\{E\}}, \overset{T_1}{\{+, *, [,], x\}}, P_1, \overset{S}{E})$, na qual:

$P_1 = \{ E \rightarrow E+E \mid E*E \mid [E] \mid x \}$

ETAPA 1: Simplificação da gramática. A gramática já está simplificada.

ETAPA 2: Transformação do lado direito das produções de comprimento maior ou igual a dois.
Exceto a produção $E \rightarrow x$, as demais devem ser substituídas como segue.

$E \rightarrow E+E \mid E*E \mid [E]$

$E \rightarrow EC_+E \mid EC_*E \mid C_[EC_]$

$C_+ \rightarrow +$

$C_* \rightarrow *$

$C_[\rightarrow [$

$C_] \rightarrow]$

$G_2 = (V_2, T_1, P_2, S)$

$V_2 = \{E, C_+, C_*, C_[, C_]\}$

$P_2 = \{ E \rightarrow EC_+E \mid EC_*E \mid C_[EC_],$

$C_+ \rightarrow +,$

$C_* \rightarrow *,$

$C_[\rightarrow [,$

$C_] \rightarrow]$

$\}$

$S = E.$

ETAPA 3: Transformação do lado direito das produções de comprimento maior ou igual a três em produções com exatamente duas variáveis.
As produções

$E \rightarrow EC_+E \mid EC_*E \mid C_[EC_]$

necessitam ser substituídas como segue:

$E \to ED_1 \mid ED_2 \mid C_[D_3$

$D_1 \to C_+ E$

$D_2 \to C_* E$

$D_3 \to E C_]$

A gramática resultante, na FNC é:

$G_{FNC} = (\{E, C_+, C_*, C_[, C_], D_1, D_2, D_3\},$
$\{+, *, [,], x\}, P_{FNC}, E)$, onde

$P_{FNC} = \{E \to ED_1 \mid ED_2 \mid C_[D_3 \mid x,$
$D_1 \to C_+ E,\ D_2 \to C_* E,\ D_3 \to E C_],$
$C_+ \to +,\ C_* \to *,\ C_[\to [,\ C_] \to]\}.$

// $\square$.

Exercícios:

exercício 6.8: Considere a seguinte gramática:

$G = (\{S, X, Y, Z, A, B\}, \{a, b, u, v\}, P, S\}$

$P = \{S \to XYZ,$

$X \to AXA \mid BXB \mid Z \mid \varepsilon,$

$Y \to AYB \mid BYA \mid Z \mid \varepsilon,$

$A \to a,$

$B \to b$

$Z \to Zu \mid Zv \mid \varepsilon\}.$

Exemplo do Sipser:

$S \to ASA \mid aB$

$A \to B \mid S$

$B \to b \mid \varepsilon$