Forme Normale de Chomsky

On dit qu'une grammaire algébrique $G = (N, \Sigma, P, S)$ est sous Forme Normale de Chomsky (FNC) si et seulement si toutes les règles sont de la forme $A \longrightarrow BC$ avec $A, B, C \in N$ ou $A \longrightarrow a$ avec $A \in N, a \in \Sigma$ ou $S \longrightarrow \varepsilon$.
La Forme Normale de Chomsky est utilisée pour démontrer certaines propriétés des grammaires algébriques. Elle permet également de construire un analyseur syntaxique par l'algorithme de Cooke, Younger et Kasami.
Théorème : Toute grammaire admet une grammaire équivalente sous FNC.

Preuve : algorithme de mise sous FNC.

Principe : On part d'une grammaire sans ε-règle (à part éventuellement la règle $S \longrightarrow \varepsilon$) ni règle unitaire (qui ne sont pas des règles sous FNC). Ainsi toutes les règles de la grammaire sont de la forme $S \longrightarrow \varepsilon$ ou $A \longrightarrow \alpha$ avec $|\alpha| \geq 1$ et telles que si $|\alpha| = 1$ alors $\alpha = a \in \Sigma$ et la règle est déjà sous FNC. Il reste donc à considérer le cas où $|\alpha| > 1$. Dans ce cas, on commence par découper les règles de telle sorte que le membre droit soit de longueur 2. On regarde ensuite si le membre droit contient un terminal a. Si tel est le cas, on remplace a par un non-terminal C_a et on ajoute si nécessaire la règle $C_a \longrightarrow a$ (qui est sous FNC).

Exemple :

$$\begin{array}{lll} S & \longrightarrow & aA|\varepsilon \\ A & \longrightarrow & ab|SbAa|bAa \end{array}$$

devient

$$\begin{array}{lll} S & \longrightarrow & C_aA|\varepsilon \\ A & \longrightarrow & C_aC_b|SD_1|C_bD_3 \\ D_1 & \longrightarrow & C_bD_2 \\ D_2 & \longrightarrow & AC_a \\ D_3 & \longrightarrow & AC_a \\ C_a & \longrightarrow & a \\ C_b & \longrightarrow & b \end{array}$$