Exemple de construction la forme normale de Chomsky d'une grammaire hors-contexte
version 1.0
le 23 mars 2009

S.V.P. ME SIGNALER TOUTE FAUTE DE FRAPPE, GRAMMAIRE, ORTOGRAPHE OU LOGIQUE; LES FAUTES DE LOGIQUE VOUS APPORTENT DES POINTS SUPPLÉMENTAIRES POUR LA NOTE FINALE. TOUTE SUGGESTION D'UNE MEILLEURE TRADUCTION D'UN TERME ANGLAIS SERA ÉGALEMENT BIENVENUE.

Soit $G = (V, \Sigma, R, S)$ la grammaire définie par
$V = \{S, T\}$
$\Sigma = \{a, b\}$
R:
$S \longrightarrow aSa \mid aTa \mid S \mid T$
$T \longrightarrow aTa \mid bTb \mid \varepsilon$

Pour la convertir en forme normale de Chomsky, on construit d'abord l'ensemble $E = \{A \in V : A \overset{*}{\Longrightarrow} \varepsilon\}$.
$E_0 = \{T\} = \{A \in V : A \longrightarrow \varepsilon\}$
$E_1 = \{T, S\} = \{A \in V : A \longrightarrow X_1 \dots X_k,\ k \geq 1,\ X_i \in E_0\} \cup E$
$E_2 = \{A \in V : A \longrightarrow X_1 \dots X_k,\ k \geq 1,\ X_i \in E_1\} \cup E_1 = E_1 = E$
La première chose à faire est d'ajouter une nouvelle variable S' avec de nouvelles règles:
$V' = V \cup \{S'\},\ S' \notin V$
$R' = (R \cup \{S' \longrightarrow \alpha : S \longrightarrow \alpha,\ \alpha \neq \varepsilon\}) =$
$\{S \longrightarrow aSa \mid aTa \mid S \mid T,$
$T \longrightarrow aTa \mid bTb \mid \varepsilon,$
$S' \longrightarrow aSa \mid aTa \mid S \mid T\}$

La règle $S' \longrightarrow \varepsilon$ sera ajoutée à la fin pour remettre ε dans le langage (car $S \in E\}$, il le faut). On enlève maintenant les règles de la forme $A \longrightarrow \varepsilon$. On remplace chaque règle $r = A \longrightarrow \alpha$ de R par un ensemble de régles obtenu en remplaçant certaines variables de α ar ε. "Certaines" veut dire "tous les choix possibles".

$S \longrightarrow aSa$ devient $\{S \longrightarrow aSa \mid aa\}$, $S \longrightarrow aTa$ devient $\{S \longrightarrow aTa \mid aa\}$, etc. pour nous donner
$R'' = (R' \cup \{S \longrightarrow aa,\ T \longrightarrow aa \mid bb,\ S' \longrightarrow aa \mid bb \mid S \mid T\}) \setminus \{T \longrightarrow \varepsilon\} =$
$\{S \longrightarrow aa \mid aSa \mid aTa \mid S \mid T,$
$T \longrightarrow aa \mid bb \mid aTa \mid bTb \mid ,$
$S' \longrightarrow aa \mid aSa \mid aTa \mid S \mid T\}$

Ensuite on enlève les règles unitaires, i.e. de la forme $A \longrightarrow B$, $A, B \in V$. On a que $S \overset{*}{\Longrightarrow} T \mid S$ et $T \overset{*}{\Longrightarrow} S \mid T$.

(enlever $S \overset{*}{\Longrightarrow} S \mid T$)
$R_{\{S\}} = \{$
$S \longrightarrow aa \mid bb \mid aSa \mid aTa \mid bTb,$
$T \longrightarrow aa \mid bb \mid aTa \mid bTb,$
$S' \longrightarrow aa \mid bb \mid aSa \mid aTa \mid T\}$

(enlever $S' \longrightarrow S \mid T$)
$R_{\{S,S'\}} = \{$
$S \longrightarrow aa \mid bb \mid aSa \mid aTa \mid bTb,$
$T \longrightarrow aa \mid bb \mid aTa \mid bTb,$
$S' \longrightarrow aa \mid bb \mid aSa \mid aTa \mid bTb, \}$

On ajoute les nouvelles variables X_x pour chaque $x \in \Sigma$:
$V_1 = \{S, T, S', X_a, X_b\}$
$R_1 = \{X_a \longrightarrow a,$
$X_b \longrightarrow b,$
$S \longrightarrow X_aX_a \mid X_aSX_a \mid X_aTX_a \mid X_bTX_b,$
$T \longrightarrow X_aX_a \mid X_bX_b \mid X_aTX_a \mid X_bTX_b,$
$S' \longrightarrow X_aX_a \mid X_aSX_a \mid X_aTX_a \mid X_bTX_b$
$\}$

On convertit chaque règle avec plus que deux variables à droite en une série de de rg̀les avec juste deux variables d̀roite; il faut ajouter des nouvelles variables. On ajoute aussi $S' \longrightarrow \varepsilon$. La grammaire finale est $\hat{G} = \hat{V}, \Sigma, \hat{R}, S'$ avec
$\hat{V} = \{S, T, S', X_a, X_b, X_S, X_T, Y_T\}$
$\hat{R} = \{$
$X_a \longrightarrow a,$
$X_b \longrightarrow b,$
$S \longrightarrow X_aX_a \mid X_aX_S \mid X_aX_T \mid X_bY_T,$
$T \longrightarrow X_aX_a \mid X_bX_b \mid X_aX_T \mid X_bY_T,$
$S' \longrightarrow X_aX_a \mid X_aX_S \mid X_aX_T \mid X_bY_T,$
$X_S \longrightarrow SX_a,$
$X_T \longrightarrow TX_a,$
$Y_T \longrightarrow TX_b,$
$S' \longrightarrow \varepsilon\}$