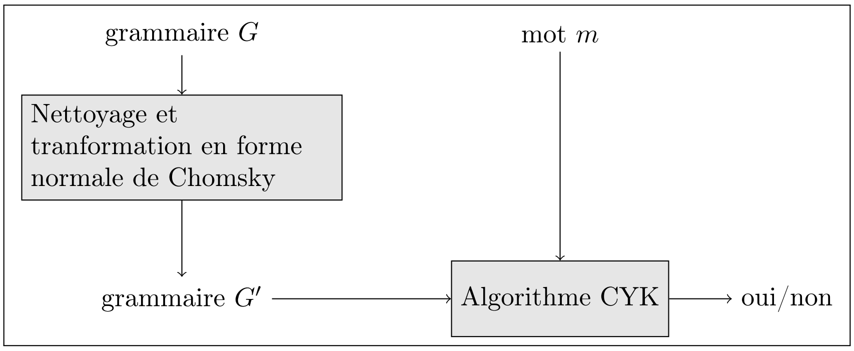
Problème du mot pour les langages algébriques

FOND1 - ENS Rennes

Problème du mot :
— Entrée : G une grammaire algébrique, w un mot ;
— Sortie : oui si $w \in L(G)$; non sinon.



15 Nettoyage de grammaires

15.1 Ne garder que les règles productives

Définition 86 (symboles productifs)**.** $\mathcal{P} = \{\alpha \in (V \cup \Sigma) \mid \exists w \in \Sigma^*, \alpha \rightarrow^* w\}$.

On garde les règles $N \rightarrow \vec{\alpha}$ où $N \in \mathcal{P}$.

Définition 87 (calcul itératif des symboles productifs)**.**
— $\mathcal{P}_0 = \Sigma$
— $\mathcal{P}_{k+1} = \mathcal{P}_k \cup \{N \in V \mid$ *il existe une règle* $N \rightarrow \alpha_1 \dots \alpha_\ell$ *où pour tout* $i \in \{1, \dots, \ell\}, \alpha_i \in \mathcal{P}_k\}$.

Proposition 88. $\mathcal{P} = \bigcup_{k \in \mathbb{N}} \mathcal{P}_k$.

15.2 Ne garder que les règles accessibles depuis l'axiome

Définition 89 (non-terminaux accessibles)**.** $\mathcal{A} = \{N \in V \mid \exists w_1, w_2 \in (\Sigma \cup V)^*, S \rightarrow^* w_1 N w_2\}$.

On garde les règles $N \rightarrow \vec{\alpha}$ où $N \in \mathcal{A}$.

Définition 90 (calcul itératif des non-terminaux accessibles)**.**
— $\mathcal{A}_0 = \{S\}$
— $\mathcal{A}_{k+1} = \mathcal{A}_k \cup \{N \in V \mid$ *il existe une règle* $M \rightarrow w_1 N w_2$ *où* $M \in \mathcal{A}_k$, $w_1, w_2 \in (\Sigma \cup V)^*\}$.

Proposition 91. $\mathcal{A} = \bigcup_{k \in \mathbb{N}} \mathcal{A}_k$.

16 Calcul des non-terminaux qui engendrent le mot vide

Définition 92 (non-terminaux qui engendrent le mot vide)**.** $\mathcal{Z} = \{N \mid N \rightarrow^* \epsilon\}$.

Définition 93 (calcul itératif de $\mathcal{Z}$)**.**
— $\mathcal{Z}_0 = \{N \in V \mid N \rightarrow \epsilon\}$
— $\mathcal{Z}_{k+1} = \mathcal{Z}_k \cup \{N \in V \mid$ *il existe une règle* $N \rightarrow \alpha_1 \dots \alpha_\ell$ *où pour tout* $i \in \{1, \dots, \ell\}$, $\alpha_i \in \mathcal{Z}_k\}$.

Proposition 94. $\mathcal{Z} = \bigcup_{k \in \mathbb{N}} \mathcal{Z}_k$.

17 Formes normales de Chomsky

17.1 Définition

Définition 95 (forme normale de Chomsky)**.** *Une grammaire est en forme normale de Chomsky si les règles sont de la forme :*

1. $N \to AB$ *où* N, A, B *sont des non-terminaux ;*
2. $N \to a$ *où* N *est un non-terminal et* $a \in \Sigma$ *;*
3. *ou éventuellement* $S \to \epsilon$.

et pas d'axiome S *dans les parties droites.*

17.2 Mise en forme normale de Chomsky

Notation 96. *On note* $|G|$ *la taille de* G *: c'est le nombre de règles plus la somme des longueurs des parties droites.*

Théorème 97. *Soit* G *une grammaire. Il existe une grammaire* G' *en forme normale de Chomsky telle que* $L(G) = L(G')$ *avec* $|G'| = O(|G|^2)$.

Étape de transformations d'une grammaire G en une grammaire G' en forme normale de Chomsky

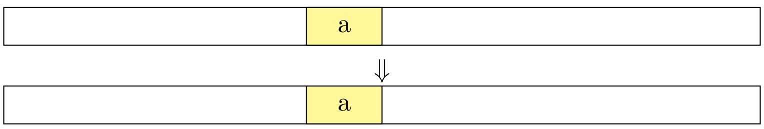
START	Ajouter $\hat{S} \to S$ où $\hat{S}$ est un nouvel axiome
BIN	Remplacer $N \to \alpha_1 \dots \alpha_n$ où $n \geq 3$ par $\begin{cases} N \to \alpha_1 X_1 \\ X_1 \to \alpha_2 X_2 \\ \vdots \\ X_{n-2} \to \alpha_{n-1}\alpha_n \end{cases}$
TERM	Si a apparaît dans une partie droite de longueur 2, — le remplacer par N_a et ajouter $N_a \to a$
DEL	Pour toutes les règles de la forme $N \to N_1 N_2$, — Si $N_1 \in \mathcal{Z}$, ajouter $N \to N_2$; — Si $N_2 \in \mathcal{Z}$, ajouter $N \to N_1$; Ajouter $\hat{S} \to \epsilon$ si $\hat{S} \in \mathcal{Z}$ Supprimer les règles $N \to \epsilon$ sauf $\hat{S} \to \epsilon$.
UNIT	Si $A \to B$ et $B \to \vec{\alpha}$, ajouter $A \to \vec{\alpha}$ Supprimer les règles $A \to B$.

18 Algorithme CYK (Cocke-Younger-Kasami)

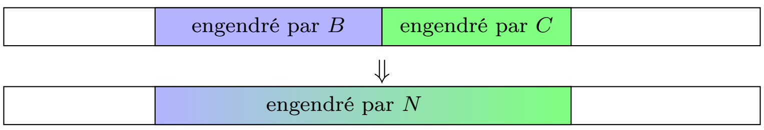
Notation 98. *Pour tout* $i, j \in \{1, \dots, |m|\}$, $i \leq j$, *on note* $E(i,j)$ *l'ensemble des non-terminaux engendrant* $m[i..j]$.

Proposition 99. $m \in L(G')$ *ssi* $S \in E(1, |m|)$.

Proposition 100.
- $E(i,i) = \{N \text{ non terminal} \mid N \to m_i\}$;
- *Si* $i < j$, $E(i,j) = \left\{ N \text{ non terminal} \mid \begin{array}{l} \text{il existe } k \in \{i, \dots, j-1\} \text{ et une règle } N \to BC \text{ dans } G' \text{ tels que} \\ B \in E(i,k) \text{ et } C \in E(k+1,j) \end{array} \right\}$.



engendré par N si $N \to a$ est une règle



si $N \to BC$ est une règle