# Module Langages Formels
# TD 7 : Lemme de l'étoile et Lemme d'Ogden

**Exercice 1** `Forme Normale de Chomsky`

**Définition**: *Grammaire CNF*
Une grammaire est sous Forme Normale de Chomsky (CNF) si toutes ses productions sont de la forme

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

**1.1.** Proposer des transformations pour les règles suivantes afin de les mettre sous forme normale de Chomsky :

1. $A \rightarrow bC$
2. $A \rightarrow Bc$
3. $A \rightarrow bc$
4. $A \rightarrow BCD$
5. $A \rightarrow bCD$
6. $A \rightarrow \alpha_1\alpha_2\alpha_3$, avec $\alpha_i \in \Sigma \cup N$
7. $A \rightarrow \alpha_1 \ldots \alpha_p$, $p \geq 3$

**1.2.** Soit $G = (\Sigma, N, P, S)$ une grammaire algébrique propre.

**1.2. 1**. En utilisant la question précédente, proposer un algorithme qui transforme $G$ en une grammaire sous forme normale de Chomsky $G'$.

**1.2. 2**. Soit $u \in \Sigma^+$. Montrer par récurrence sur $n$ que $\forall A \in N,\ A \vdash_G^n u \Rightarrow A \vdash_{G'}^* u$.

**1.2. 3**. Soit $u \in \Sigma^+$. Montrer par récurrence sur $n$ que $\forall A \in N,\ A \vdash_{G'}^n u \Rightarrow A \vdash_G^* u$.

Ainsi, toute grammaire algébrique propre admet une grammaire sous forme normale de Chomsky équivalente.

**Exercice 2** `Lemme d'Ogden`

L'objectif de cet exercice est de démontrer le Lemme suivant :

**Lemme (Ogden)** :
Soit $L$ un langage algébrique. Il existe un entier $N$ tel que pour tout mot $z \in L$ dans lequel on marque au moins $N$ positions distinctes, il est possible de décomposer $z$ sous la forme $z = uxvyw$ avec
- $x$ ou $y$ contient au moins une position marquée,
- $xvy$ contient au plus $N$ positions marquées,
- pour tout $i \geq 0$, $ux^ivy^iw \in L$.

**Définition**:
Soit $T$ un sous-arbre d'un arbre de dérivation selon une grammaire CNF. On suppose marquées certaines feuilles de $T$. On appelle *embranchement* un nœud de $T$ ayant deux fils, tel que chacun de ses fils contienne au moins une feuille marquée.

**2.1.** Soit $T$ un sous-arbre d'un arbre de dérivation selon une grammaire CNF. On suppose qu'au moins $2^h$ feuilles distinctes de $T$ ont été marquées.

Montrer qu'il existe un chemin, d'une feuille à la racine, passant par au moins $h$ embranchements et tel que pour tout $i$, le $i$-ème embranchement ait au plus $2^i$ descendants marqués.

**2.2.** On considère une grammaire CNF $G$ engendrant le langage $L$. Montrer qu'il existe un entier $N$ tel que :
- pour tout mot $w \in L$ dans lequel on marque au moins $N$ positions,
- pour tout arbre de dérivation de $w$,

il existe deux embranchements $b_1$ et $b_2$ tels que
- $b_1$ est un ancêtre de $b_2$,
- $b_1$ est un ancêtre d'au plus $N$ feuilles marquées,
- $b_1$ et $b_2$ sont étiquetés par la même variable.

**2.3.** En déduire le lemme d'Ogden.

**2.4.** On s'intéresse au langage $L = \{a^ib^jc^kd^l \quad | \quad i = 0 \quad \text{ou} \quad j = k = l\}$.

**2.4. 1.** Montrer que pour tout $N \in \mathbb{N}^+$ et tout mot $z \in L$ avec $|z| \geq N$, il existe une décomposition $z = uxvyw$ telle que
- $|xy| \geq 1$
- $|xvy| \leq N$
- pour tout $i \geq 0$, $ux^ivy^iw \in L$.

**2.4. 2.** Montrer que $L$ n'est pas algébrique.