

TD8. Lemme d'Ogden

L'objectif de ce TD est de montrer une version plus forte du lemme de l'étoile pour les langages algébriques :

Lemme 1 (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 $z = uxvtyw$ avec

- x ou y contient au moins une position marquée,
- xvy contient au plus N positions marquées,
- pour tout $i \geq 0$, $ux^i v y^i w \in L$.

Question 1. On se donne une grammaire algébrique G engendrant un langage L . Montrer qu'il existe une grammaire sous forme normale de Chomsky (CNF) reconnaissant le langage $L - \{\varepsilon\}$.

Rappel : Une grammaire CNF est une grammaire où les productions sont toutes de la forme

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

Définition 1. 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.

Question 2. 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.

Question 3. 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.

Question 4. En déduire le lemme d'Ogden.

Question 5. On s'intéresse au langage $L_3 = \{a^i b^j c^k d^l \mid i = 0 \text{ ou } j = k = l\}$.

Montrer que pour tout $N \in \mathbb{N}$ et tout mot $z \in L_3$ il existe une décomposition $z = uxvtyw$ telle que

- $|xy| \geq 1$
- $|xvy| \leq N$
- pour tout $i \geq 0$, $ux^i v y^i w \in L_3$.

Montrer que L_3 n'est pas algébrique.