

Exercices Mots et Langages

1 A traiter après lecture de 1.1,1.2 et 1.3

1. Déterminer tous les mots dont tous les facteurs sont des préfixes.
2. Les mots de Fibonacci sur l'alphabet $\Sigma = \{a, b\}$ sont définis par récurrence par $f_0 = \epsilon, f_1 = a, f_2 = b$. et, pour tout $n \geq 3, f_n = f_{n-1}f_{n-2}$.
 - (a) Montrer que, pour $n \geq 3$, le suffixe de longueur 2 de f_n est ab si n est pair, ba si n est impair.
 - (b) Pour $n \geq 3$, notons g_n le mot déduit de f_n en enlevant le suffixe de longueur 2 précisé à la question précédente. Montrer que g_n est un palindrome.
3. Prouver le théorème 2 de la partie 1.4 par récurrence sur $|xy|$.
4. Soient u et v des mots non vides. Montrer qu'il existe des mots x et y tels que $u = xy$ et $v = yx$ si, et seulement s'il existe un mot w tel que $uw = wv$.

2 TD lundi 21/09

1. On note pour un langage L , $\text{pref}(L)$ l'ensemble de ses préfixes. Exprimer $\text{pref}(L_1L_2)$ en fonction de $\text{pref}(L_1), \text{pref}(L_2), L_1$ et L_2 .
2. Soit L un langage non vide. Montrer que $\epsilon \in L$ ssi $L \subset LL$ (on pourra considérer u le mot de L de longueur minimale).
3. Montrer que la concaténation est distributive à gauche et à droite sur l'union.
4. Le lemme d'Arden est un résultat extrêmement classique. L'objectif ici est d'étudier des (systèmes) d'équations dont les inconnues sont des langages.

Dans cette partie, Σ est un alphabet et $K, L \in \mathcal{P}(\Sigma^*)$ sont deux langages sur cet alphabet. On y étudie l'équation $(E) : X = KX + L$ dont l'inconnue est le langage X .

- (a) Montrer que le langage K^*L est solution de l'équation (E) .
 - (b) Montrer que toute solution X de (E) vérifie $K^*L \subset X$.
 - (c) On suppose ici que $\epsilon \notin K$. Montrer sous cette condition que toute solution X de (E) vérifie $X \subset K^*L$.
*Indication : On pourra procéder par l'absurde et considérer un mot de longueur minimale de $X \setminus K^*L$.*
 - (d) Dédire des questions précédentes le lemme d'Arden : Pour tout alphabet Σ , pour tous $L, K \in \mathcal{P}(\Sigma^*)$ tels que $\epsilon \notin K$, l'équation $X = KX + L$ admet une unique solution, à savoir K^*L .
 - (e) A-t-on toujours unicité de la solution à l'équation (E) si $\epsilon \in K$? Justifier.
5. Dans cette partie, $\Sigma = \{a, b\}$. On note L_1 le langage des mots ayant un nombre pair de b et L_2 le langage des mots ayant un nombre impair de b .

- (a) Expliquer pourquoi (L_1, L_2) est solution du système d'équations suivant :

$$(S) : \begin{cases} L_1 = aL_1 + bL_2 + \epsilon & (1) \\ L_2 = aL_2 + bL_1 & (2) \end{cases}$$

- (b) En utilisant le lemme d'Arden, résoudre le système (S) et en déduire une expression rationnelle pour L_1 et L_2 . *Indication : utiliser le lemme d'Arden sur (2) puis substituer dans (1).*

3 DM pour le mardi 22/09

On pourra utiliser le théorème d'Euler suivant : Dans un graphe orienté $G = (V, E)$ d'ordre au moins 2, un cycle eulérien est un cycle qui passe une fois et une seule par chaque arête. G est dit eulérien s'il admet un cycle eulérien. Le théorème d'Euler affirme qu'un graphe G fortement connexe est eulérien si, et seulement si, le degré entrant de chaque sommet est égal à son degré sortant.

Soit Σ un alphabet et $u = u_1 \dots u_n \in \Sigma^*$ un mot non vide. Pour $i > n$, on note u_i la lettre $u_{i'}$ tel que $i \equiv i' \pmod n$ et $i' \in \{1, \dots, n\}$.

On appelle facteur circulaire de u un mot $v = v_1 \dots v_p \in \Sigma^*$ tel que $p \leq n$ et il existe $k \in \{0, \dots, n-1\}$ tel que $v_i = u_{i+k}$.

Enfin, pour $k \in \mathbb{N}^*$, on appelle mot de de Bruijn d'ordre k un mot u tel que tous les mots de Σ^k apparaissent une unique fois comme facteur circulaire de u .

Par exemple, $aabb$ est un mot de de Bruijn d'ordre 2 et $aaababbb$ est un mot de de Bruijn d'ordre 3 pour $\Sigma = \{a, b\}$.

Le but de cet exercice est de construire une suite de mots $(b_k)_{k \in \mathbb{N}^*}$ telle que pour tout k , b_k est un mot de de Bruijn d'ordre k . On considère pour la suite un alphabet Σ quelconque.

1. Déterminer la taille de b_k s'il existe.
2. Quel mot peut convenir pour b_1 ?

On propose la construction suivante : Soit $G_k = (V_k, E_k)$ le graphe orienté tel que :

- $V_k = \Sigma^k$ (chaque mot de longueur k est un sommet) ;
 - E_k est défini de la façon suivante : il existe une arête (u, v) si, et seulement si, le suffixe de u de longueur $k-1$ est préfixe de v (y compris dans les éventuels cas où $u = v$).
3. Construire G_2 pour $\Sigma = \{0, 1\}$.
 4. En utilisant le théorème d'Euler ci-dessus, montrer qu'il existe un cycle dans G_k empruntant une et une seule fois chaque arc.
 5. En étiquetant les arêtes judicieusement, en déduire une façon de construire b_{k+1} à partir du graphe G_k .
 6. Déterminer un mot de de Bruijn d'ordre 4 sur $\Sigma = \{a, b\}$ et un mot de de Bruijn d'ordre 2 sur $\Sigma = \{a, b, c\}$.