# Problème 9

Soit $A$ un ensemble fini. Un mot fini sur $A$ est une suite finie $u = (u_n)_{1\leqslant n\leqslant l}$ d'éléments de $A$ ; on le note plus communément par concaténation $u = u_1 \dots u_l$. L'entier $l$ est la longueur de $u$ ; on la note $|u|$. Un mot infini sur $A$ est une suite $\mathbf{a} = (a_n)_{n\geqslant 1}$ avec $a_n \in A$. Un mot fini $u$ est un facteur d'un mot (fini ou infini) $v$ s'il existe un entier $j$ tel que $u = v_j \dots v_{j+|u|-1}$.

Un mot infini $\mathbf{a}$ est à lacunes bornées si pour tout facteur $u$ de $\mathbf{a}$, il existe un entier $l$ tel que pour tout facteur $v$ de longueur $l$ de $\mathbf{a}$, $u$ est un facteur de $v$.

Un mot infini $\mathbf{a}$ est un $T$-mot s'il satisfait à la propriété suivante

$$\forall n \in \mathbb{N}, \exists p \in \mathbb{N}\setminus\{0\}, \forall k \in \mathbb{N}, a_{n+kp} = a_p.$$

Soient $\mathbf{a}, \mathbf{b}$ deux mots infinis sur des alphabets $A$ et $B$ respectivement. Le produit de $\mathbf{a}$ et $\mathbf{b}$ est le mot infini $\mathbf{c}$ sur l'alphabet $A \times B$ tel que $c_n = \begin{pmatrix} a_n \\ b_n \end{pmatrix}$.

1. Montrer que tout $T$-mot est à lacunes bornées.
2. Soit $\mathbf{d}$ le mot infini défini par

$$d_n = \begin{cases} 1 & \text{si} \quad s_2(n) \neq s_2(n+1) \mod 2 \\ 0 & \text{sinon} \end{cases}$$

   où $s_2(n)$ est la somme des chiffres de la représentation de $n$ en base 2. Montrer que $\mathbf{d}$ est un $T$-mot.
3. Montrer que si $\mathbf{a} \in A^{\mathbb{N}}$ est un $T$-mot et si $\mathbf{b} \in B^{\mathbb{N}}$ est à lacunes bornées, alors $\mathbf{a} \times \mathbf{b}$ est à lacunes bornées.

**Solution des auteurs :** 1. Soit $\mathbf{a} = (a_n)_{n\in\mathbb{N}} \in A^{\mathbb{N}}$ un $T$-mot et soit $u = u_1 \cdots u_\ell$ un facteur de $\mathbf{a}$. Il existe $n \in \mathbb{N}$ tel que $a_n a_{n+1} \cdots a_{n+\ell-1} = u$.

Le mot $\mathbf{a}$ étant un $T$-mot, pour tout $i \in \{n, n+1, \dots, n+\ell-1\}$, il existe $p_i \in \mathbb{N}\setminus\{0\}$ tel que

$$a_{i+kp_i} = a_i \qquad \text{pour tout } k \in \mathbb{N}.$$

On considère $p = \operatorname{ppcm}(\{p_n, p_{n+1}, \dots, p_{n+\ell-1}\})$. Pour tout $k \in \mathbb{N}$, on a donc

$$a_{n+kp} a_{n+kp+1} \cdots a_{n+kp+\ell-1} = u.$$

Par conséquent, pour $L = \max\{p, n\} + \ell$, tout facteur de longueur $L$ de $\mathbf{a}$ contient une occurrence du mot $u$.

2. Pour $n \in \mathbb{N}$, notons $\operatorname{rep}_2(n)$ le mot $a_\ell a_{\ell-1} \cdots a_0$ sur l'alphabet $\{0,1\}$ tel que $a_\ell \neq 0$ et

$$n = \sum_{i=0}^{\ell} a_i 2^i.$$

Notons également $\varepsilon$ l'unique mot de longueur $0$ ; on a $\operatorname{rep}_2(0) = \varepsilon$.

Soit $n \in \mathbb{N}$. Notons $k$ le plus grand entier tel que $1^k$ est suffixe de $\operatorname{rep}_2(n)$. On a donc $\operatorname{rep}_2(n) = 1^k$ ou $\operatorname{rep}_2(n) = u01^k$ pour un mot $u$ sur $\{0,1\}$. On a alors, respectivement, $\operatorname{rep}_2(n+1) = 10^k$ et $\operatorname{rep}_2(n+1) = u10^k$. Par conséquent, on obtient

$$d_n = \begin{cases} 1, & \text{si } k \text{ est pair ;} \\ 0, & \text{sinon.} \end{cases}$$

Prenons alors $p = 2^{k+1}$, de sorte que $\operatorname{rep}_2(p) = 10^{k+1}$ et, pour tout entier $m > 0$, $0^{k+1}$ est suffixe de $\operatorname{rep}_2(mp)$. Pour tout $m \in \mathbb{N}$, il existe alors un mot $v$ sur $\{0,1\}$ tel que $\operatorname{rep}_2(n+mp) = v10^k$. On en déduit $d_{n+mp} = d_n$ pour tout $m \in \mathbb{N}$.

3. Soit $u = (v, w)$ un facteur de longueur $\ell$ de $\mathbf{a} \times \mathbf{b}$. Soit $N \in \mathbb{N}$ la première position de $u$ dans $\mathbf{a} \times \mathbf{b}$, i.e., $N$ est le plus petit entier tel que

$$a_N a_{N+1} \cdots a_{N+\ell-1} = v \quad \text{et} \quad b_N b_{N+1} \cdots b_{N+\ell-1} = w.$$

Le mot $\mathbf{b}$ étant un $T$-mot, les mêmes arguments que pour le point 1 montrent qu'il existe une constante $p > 0$ telle que pour tout $n \in \mathbb{N}$,

$$b_{N+np} b_{N+np+1} \cdots b_{N+np+\ell-1} = w.$$

On va montrer l'existence d'une constante $K$ telle que pour tout $n \in \mathbb{N}$, si $v$ apparaît dans $\mathbf{a}$ à la position $N + np$, alors il existe $0 < n' < K$ tel que $v$ apparaît encore à une position $N + np + n'p$. Dans ce cas, le facteur $u$ apparaît dans tout facteur de $\mathbf{a} \times \mathbf{b}$ de longueur $|u| + \max\{N, Kp\}$.

On définit le mot infini $\mathbf{d} = (d_n)_{n \in \mathbb{N}}$ sur $\{0, 1\}$ par

$$d_n = \begin{cases} 1, & \text{si } a_n \cdots a_{n+\ell-1} = v; \\ 0, & \text{sinon.} \end{cases}$$

On a $d_N = 1$ et, le mot $\mathbf{a}$ étant à lacunes bornées, le mot $\mathbf{d}$ l'est également.

Considérons un entier $n \in \mathbb{N}$ tel que $d_{N+np} = 1$. Le mot $\mathbf{d}$ étant à lacunes bornées, il existe une constante $K(p)$ telle que tout facteur de $\mathbf{d}$ de longueur $p$ apparaît dans tout facteur de $\mathbf{d}$ de longueur $K(p)$. Par conséquent, le facteur $d_{[N+np, N+np+p-1]}$ apparaît au moins deux fois dans le facteur $d_{[N+np, N+np+K(p)]}$. En itérant cet argument, on construit une suite de mots $u_0, u_1, \ldots, u_p$ tels que

— $u_0 = d_{[N+np, N+np+p-1]}$ ;

— pour tout $i \geqslant 1$, $u_i$ est le plus petit préfixe de $(d_{N+np+m})_{m \in \mathbb{N}}$ qui contient deux occurrences de $u_{i-1}$.

On définit également une suite de constantes $K_0, K_1, \ldots, K_p$ telles que

— $K_0 = p - 1$ ;

— pour tout $i \geqslant 1$, $K_i$ est le plus petit entier tel que tout facteur de longueur $1 + K_{i-1}$ de $\mathbf{d}$ apparaît dans tout facteur de longueur $K_i$ de $\mathbf{d}$.

Observons que la suite de constantes $K_0, K_1, \ldots, K_p$ ne dépend pas de la suite de mots $u_0, u_1, \ldots, u_p$, mais uniquement de $p$. Autrement dit, la constante $K_p$ ne dépend pas de la position $N + np$ initialement choisie. De plus, pour tout $i \in \{0, 1, 2 \ldots, p\}$, le facteur $u_i$ est de longueur au plus $1 + K_i$. En effet, le préfixe de longueur $1 + K_i$ de $(d_{N+np+m})_{m \in \mathbb{N}}$ contient au moins deux occurrences de $u_{i-1}$ : une occurrence comme préfixe et une occurrence dans le facteur $d_{[N+np+1, N+np+K_i]}$.

Pour tous $p \geqslant j \geqslant i \geqslant 0$, notons $m(i, j)$ le plus grand entier tel que

$$d_{[N+np+m(i,j), N+np+m(i,j)+|u_i|-1]} = u_i \quad \text{et} \quad m(i, j) + |u_i| \leqslant |u_j|.$$

Autrement dit, $m(i, j)$ est la dernière position de $u_i$ comme facteur du préfixe $u_j$ de $(d_{N+np+m})_{m \in \mathbb{N}}$. Le mot $u_i$ étant suffixe de $u_j$ pour tous $i \leqslant j$, on a, pour tous $k \leqslant i \leqslant j$,

$$m(k, j) = m(k, i) + m(i, j).$$

Par le principe des tiroirs, il existe $0 \leqslant I < J \leqslant p$ tels que

$$m(0, I) \equiv m(0, J) \bmod p.$$

Par conséquent, on a, pour tout $k \leqslant I$,

$$m(k, I) = m(0, I) - m(0, k) \quad \text{et} \quad m(k, J) = m(0, J) - m(0, k),$$

et donc

$$m(k, I) \equiv m(k, J) \bmod p.$$

En particulier, pour $k = I$, on a

$$0 = m(I, I) \equiv m(I, J) \bmod p.$$

Comme $J > 0$, on a $m(I,J) = n'p$ pour un entier $n' > 0$. Par définition, le mot $u_I$, apparaît dans $\mathbf{d}$ à la position $N + np + n'p$ avec $n'p \leqslant K_p$.

On a donc montré que la constante $K = K_p/p$ est telle que pour tout $n \in \mathbb{N}$, si $d_N + np = 1$, alors il existe un entier $0 < n' \leqslant K$ tel que $d_{N+np+n'p} = 1$, ce qui conclut la preuve.

**Solution de l'équipe Team Parc :**

## Question 1

Soit **a** un $T$-mot, et $u = a_n a_{n+1} \cdots a_{n+r}$ un facteur de **a**. Pour chaque $i \in [\![0, r]\!]$, il existe un entier strictement positif $p_i$ tel que, pour tout $k_i \in \mathbb{N}$, on ait $a_{n+i+k_i p_i} = a_{n+i}$. Posons $p = \prod_{i=0}^{r} p_i$. En particulier, en choisissant $k_i = c\frac{p}{p_i}$, où $c$ est un entier positif arbitraire, on remarque que $a_{n+i+cp} = a_{n+i}$ pour tout $i \in [\![0, r]\!]$.

Posons $l = n + p + r$. Soit $v$ un facteur de **a** de longueur $l$, et montrons que $u$ est un facteur de $v$. Tout d'abord, commençons par ignorer les $n$ premières lettres de $v$, sur lesquelles on peut potentiellement n'avoir aucune information. On dispose d'un facteur de $v$ de la forme $a_{n+k_0} \cdots a_{n+k_0+p+r-1}$ pour un certain entier positif $k_0$. Parmi les $p$ entiers $\{k_0, k_0+1, \cdots, k_0+p-1\}$, notons $k_1$ celui qui est divisible par $p$. De cette manière, $a_{n+k_1} \cdots a_{n+r+k_1}$ est un facteur de $v$. En notant $c$ l'entier positif tel que $k_1 = cp$, on remarque que :

$$a_{n+k_1} \cdots a_{n+r+k_1} = a_{n+cp} \cdots a_{n+r+cp} = a_n \cdots a_{n+r} = u.$$

En définitive, $u$ est un facteur de $v$, ce qui prouve que **a** est à lacunes bornées.

## Question 2

Montrons tout d'abord que $d_n = 1$ si et seulement si la représentation binaire de $n$ se termine par un nombre pair de 1. Soit $n \in \mathbb{N}$ et $\overline{c_d c_{d-1} \cdots c_0}^2$ son écriture en base 2. On suppose qu'au moins un des chiffres est zéro, quitte à en rajouter un au début du nombre. Notons $r$ l'indice du dernier zéro, de telle sorte que :

$$n = \overline{c_d \cdots c_{r+1} 0 \underbrace{1 \cdots 1}_{\text{r chiffres "1"}}}^2.$$

On remarque que $n+1$ s'écrit :

$$n + 1 = \overline{c_d \cdots c_{r+1} 1 \underbrace{0 \cdots 0}_{\text{r chiffres "0"}}}^2.$$

Il est alors clair que les sommes des chiffres de $n$ et $n+1$ écrits en base 2 sont distinctes modulo 2 si et seulement si $r$ est pair.

A présent, montrons que **d** est un $T$-mot. Soit $n \in \mathbb{N}$, et posons $p = 2(n+1)$. On va prouver que $d_{n+kp} = d_n$ pour $k \in \mathbb{N}$. Sans prendre en compte les chiffres situés avant le dernier zéro, on écrit :

$$n = \overline{\cdots 0 \underbrace{1 \cdots 1}_{\text{r chiffres "1"}}}^2.$$

On constate alors que, pour $k \in \mathbb{N}$ :

$$\begin{aligned}
n + kp = n + 2k(n+1) &= \overline{\cdots 0 \underbrace{1 \cdots 1}_{\text{r chiffres "1"}}}^2 + 2k \times \overline{\cdots 1 \underbrace{0 \cdots 0}_{\text{r chiffres "0"}}}^2 \\
&= \overline{\cdots 0 \underbrace{1 \cdots 1}_{\text{r chiffres "1"}}}^2 + k \times \overline{\cdots \underbrace{1 0 \cdots 0}_{\text{r+1 chiffres "0"}}}^2 \\
&= \overline{\cdots 0 \underbrace{1 \cdots 1}_{\text{r chiffres "1"}}}^2 + \overline{\cdots \underbrace{0 \cdots 0}_{\text{r+1 chiffres "0"}}}^2 \\
&= \overline{\cdots 0 \underbrace{1 \cdots 1}_{\text{r chiffres "1"}}}^2.
\end{aligned}$$

Ainsi, le nombre de 1 à la fin de l'écriture binaire est le même pour $n$ et pour $n + kp$, ce qui prouve que $d_{n+kp} = d_n$, et donc que $\mathbf{d}$ est un $T$-mot.

## Question 3

Soit $\mathbf{a} \in A^{\mathbb{N}}$ un $T$-mot et $\mathbf{b} \in B^{\mathbb{N}}$ à lacunes bornées. Soit $u = (a_n, b_n)(a_{n+1}, b_{n+1}) \cdots (a_{n+r}, b_{n+r})$ un facteur de $\mathbf{a} \times \mathbf{b}$. On a prouvé à la question 1 l'existence d'un entier strictement positif $p$ qui vérifie, pour tout entier naturel $k$ :

$$a_{n+kp} \cdots a_{n+r+kp} = a_n \cdots a_{n+r}.$$

Notons $P$ l'ensemble des classes $\omega \in \mathbb{Z}/p\mathbb{Z}$ pour lesquelles il existe un indice $i \geqslant n$ dont la classe est $\omega$ et qui vérifie $b_i \cdots b_{i+r} = b_n \cdots b_{n+r}$. En particulier, $\overline{n} \in P$. Notons $m$ le cardinal de $P$. On choisit $m$ indices de classes distinctes $n = i_0 < \cdots < i_{m-1}$ tels que, pour chaque $j \in [\![0, m-1]\!]$ :

$$b_{i_j} \cdots b_{i_j+r} = b_n \cdots b_{n+r}.$$

De cette façon, $\{\overline{i_0}, \cdots, \overline{i_{m-1}}\} = P$.

Puisque $\mathbf{b}$ est à lacunes bornées, il existe $l_0 \in \mathbb{N}$ tel que tout facteur de longueur $l_0$ de $\mathbf{b}$ admette $b_{i_0} b_{i_0+1} \cdots b_{i_{m-1}+r}$ comme facteur. Posons $l = n + l_0$ et considérons un facteur $v$ de longueur $l$ de $\mathbf{a} \times \mathbf{b}$. Nous allons montrer que $u$ est un facteur de $v$. En premier lieu, ignorons les $n$ premières lettres de $v$. On obtient un facteur de $v$ de longueur $l_0$. Il existe donc un entier positif $d$ tel que :

$$b_{i_0+d} \cdots b_{i_{m-1}+r+d} = b_{i_0} \cdots b_{i_{m-1}+r},$$

où $(a_{i_0+d}, b_{i_0+d}) \cdots (a_{i_{m-1}+r+d}, b_{i_{m-1}+r+d})$ est un facteur de $v$.

Remarquons que $\{\overline{i_0 + d}, \overline{i_1 + d}, \cdots, \overline{i_{m-1} + d}\} = P$. En effet, l'égalité ci-dessus indique que, pour $j \in [\![0, m-1]\!]$ :

$$b_{i_j+d} \cdots b_{i_j+r+d} = b_{i_j} \cdots b_{i_j+r},$$

qui est égal, par définition, à $b_n \cdots b_{n+r}$. Ainsi, $\overline{i_j + d} \in P$, ce qui montre que $\{\overline{i_0 + d}, \overline{i_1 + d}, \cdots, \overline{i_{m-1} + d}\} \subset P$. L'inclusion réciproque est alors automatique par cardinalité.

Puisque $\overline{n} \in P$, on déduit de l'égalité précédente qu'il existe $j \in [\![0, m-1]\!]$ tel que $\overline{i_j + d} = \overline{n}$. Notons $k$ l'entier positif vérifiant $i_j + d = n + kp$. On remarque alors que, d'une part :

$$a_{i_j+d} \cdots a_{i_j+r+d} = a_{n+kp} \cdots a_{n+r+kp} = a_n \cdots a_{n+r},$$

et que d'autre part :

$$b_{i_j+d} \cdots b_{i_j+r+d} = b_n \cdots b_{n+r}.$$

En définitive, on trouve :

$$(a_{i_j+d}, b_{i_j+d}) \cdots (a_{i_j+r+d}, b_{i_j+r+d}) = u,$$

où $(a_{i_j+d}, b_{i_j+d}) \cdots (a_{i_j+r+d}, b_{i_j+r+d})$ est un facteur de $v$. Cela prouve que $\mathbf{a} \times \mathbf{b}$ est à lacunes bornées.