ECOLES NORMALES SUPERIEURES

CONCOURS D'ADMISSION 2025

VENDREDI 19 DECEMBRE 2025
14h00 - 20h00

FILIERE MPI - Epreuve n° 8

INFORMATIQUE D (L)

Durée: 6 heures

L'utillisation des calculatrices n'est pas autorisée pour cette épreuve

Le sujet comprend 7 pages, numérotées de 1 à 7

Espaces de shift et mots sturmien

* * *

On cherche dans ce sujet à étudier les mots et morphismes sturmiens. Un mot infini est dit sturmien s'il possède $n + 1$ facteurs distincts de longueur n. On montrera le théorème de **Mignosi et Séébold** qui caractérise les morphismes sturmiens-préservant comme étant exactement la composition de 3 morphismes de base.

* * *

Le problème comprend une partie préliminaire sur les espaces de shift, puis 4 parties numérotées de 1 à 4. La partie 1 est indépendante des autres et fait le lien entre espaces de shift et langages réguliers. La partie 2 introduit les notions de morphisme et substitution, importantes aux parties 3 et 4. La partie 3 introduit les mots sturmiens et la partie 4 traite de la démonstration du théorème de Mignosi et Séébold.

À tout moment il est possible d'admettre le résultat d'une question et de l'utiliser ultérieurement, à condition de l'indiquer clairement.

Le sujet est long et comporte des questions délicates : il est tout à fait normal de bloquer sur des questions et il est possible de réussir l'épreuve sans traiter une part substantielle du sujet.

La clarté, la concision et la précision de la rédaction seront prises en compte dans la notation.

* * *

Notations

On note $\mathbb{N} = \{0, 1, \ldots\}$ l'ensemble des naturels et $\mathbb{N}^* = \{1, 2, \ldots\}$ l'ensemble des naturels strictements positifs. On note $\mathbb{R}$ l'ensemble des nombres réels et $\mathbb{R}_+$ l'ensemble des réels positifs ou nuls. Pour A un ensemble on note $|A|$ son cardinal. Pour A et B deux ensembles on note B^A l'ensemble des applications de A vers B. Pour un ensemble A, on note $\mathcal{P}(A)$ l'ensemble de ses parties. Une suite $(u_n)_{n\in\mathbb{N}} \in A^{\mathbb{N}}$ est dite *ultimement périodique* si $\exists N, T \in \mathbb{N}, \forall n > N, u_n = u_{n+T}$

Dans tout le sujet, Σ sera un alphabet fixé à au moins 2 lettres. L'étoile de Kleene de Σ est noté Σ^*. Soit $u \in \Sigma^*$, on note $|u|$ sa longeur et pour $\alpha \in \Sigma$, on note $|u|_\alpha$ le nombre d'occurences de α dans u. Soient $u, v \in \Sigma^*$ on écrira uv leur concaténation. Si $u \in \Sigma^*$ est un mot, on pourra écrire $u = (u_0, u_1, \ldots, u_{n-1})$ ou $u = u_0 u_1 \ldots u_{n-1}$, avec $n = |u|$. Pour $u = u_0 \ldots u_{|u|-1}$ un mot, on note $\overline{u} = u_{|u|-1} \ldots u_0$ le *miroir* de u. On dit que $u \in \Sigma^*$ est un *palindrome* si $u = \overline{u}$

Un *automate* $\mathcal{A}$ sera représenté par un quintuplet $\mathcal{A} = \langle \Sigma, Q, I, \delta, F \rangle$ avec $F, I \subseteq Q$ et $\delta \subseteq Q \times \Sigma \times Q$. Pour $\mathcal{A}$ un automate, on définit $L(\mathcal{A})$ le langage qu'il reconnait par l'ensemble des $w \in \Sigma^*$ tel qu'il existe $q_0, \ldots, q_{|u|}$ avec $q_0 \in I$, $q_{|u|} \in F$ et $\forall i < |u|, (q_i, u_i, q_{i+1}) \in \delta$. Un automate sera dit *déterministe* si $|I| = 1$ et que pour tout $(q, \alpha) \in Q \times \Sigma$ on a $|\{q' \in Q \mid (q, \alpha, q') \in \delta\}| \leq 1$. Il sera dit *complet* si pour tout $(q, \alpha) \in Q \times \Sigma$ on a $|\{q' \in Q \mid (q, \alpha, q') \in \delta\}| \geq 1$. Un automate n'est pas forcément complet ou déterministe.

* * *

Partie préliminaire

On appelle *mot infini* tout élément de $\Sigma^{\mathbb{N}}$. Pour $u_0..u_{n-1} \in \Sigma^*$ et $v \in \Sigma^{\mathbb{N}}$, on définit uv la concaténation de u et v par $(u_0, \ldots, u_{n-1}, v_0, v_1, \ldots) \in \Sigma^{\mathbb{N}}$. On définit $\sigma : \Sigma^{\mathbb{N}} \to \Sigma^{\mathbb{N}}$ par $\sigma((u_0, u_1, \ldots)) := (u_1, u_2, \ldots)$ (on retire la première lettre à u). σ sera appelé *shift*. Pour $u, v \in \Sigma^{\mathbb{N}}$ deux mots infini, on définie la distance de u à v par :

$$d(u,v) := \begin{cases} 0 & \text{si } u = v \\ |\Sigma|^{-\min\{i \in \mathbb{N} \mid u_i \neq v_i\}} & \text{sinon} \end{cases}$$

(1) Montrer les 3 propriétés suivantes sur d (on dira alors que d est une distance) :
- Montrer que $\forall x, y \in \Sigma^*, d(x,y) = d(y,x)$
- Montrer que $\forall x, y \in \Sigma^*, d(x,y) = 0 \Leftrightarrow x = y$
- Montrer que $\forall x, y, z \in \Sigma^*, d(x,z) \leq d(x,y) + d(y,z)$

Pour $u \in \Sigma^{\mathbb{N}}$, on définit $\mathcal{O}(u) = \{\sigma^k(u) : k \in \mathbb{N}\}$. Soit $(u_i)_{i \in \mathbb{N}} \in (\Sigma^{\mathbb{N}})^{\mathbb{N}}$ une suite de mots infinis, on dit que $u_n \underset{n \to \infty}{\longrightarrow} v$ si $d(u_n, v) \underset{n \to \infty}{\longrightarrow} 0$. Pour $X \subseteq \Sigma^{\mathbb{N}}$, on définit $\overline{X}$ l'ensemble des $v \in \Sigma^{\mathbb{N}}$ tel qu'il existe $(u_i)_{i \in \mathbb{N}} \in X^{\mathbb{N}}$ tel que $u_n \underset{n \to \infty}{\longrightarrow} v$. Autrement dit, $\overline{X}$ est *l'adhérence* de X pour la distance d.

On dit que $X \subseteq \Sigma^{\mathbb{N}}$ est un *espace de shift* si $\sigma(X) \subseteq X$ et $X = \overline{X}$.

(2) Montrer que si $u_n \underset{n \to \infty}{\longrightarrow} v$ et $u_n \underset{n \to \infty}{\longrightarrow} v'$, alors $v = v'$.

On note alors, sous condition d'existence, $\lim_{n \to \infty} u_n$ l'unique v tel que $u_n \underset{n \to \infty}{\longrightarrow} v$.

(3) Montrer que pour tout $u \in \Sigma^{\mathbb{N}}$, $\overline{\mathcal{O}(u)}$ est un espace de shift.

(4) Montrer que $\overline{\mathcal{O}(u)}$ est fini si et seulement si $u \in \Sigma^{\mathbb{N}}$ est ultimement périodique.

Pour X un espace de shift, on définit $F(X) = \{u \in \Sigma^* \mid \exists v \in \Sigma^{\mathbb{N}}, uv \in X\}$.

(5) Montrer que $X \mapsto F(X)$ est injective.

* * *

1 Espace sofique et type fini

On dit qu'un espace de shift X est *sofique* si $F(X)$ est un langage régulier.

(6) Montrer que si X est sofique alors il existe un automate $\mathcal{A} = \langle \Sigma, Q, I, \delta, F' \rangle$ qui reconnait $F(X)$ avec $Q = I = F'$.

Dans cette partie, **tous** les automates seront supposés tels que $Q = F' = I$. On définira donc un automate seulement par le triplet $\langle \Sigma, Q, \delta \rangle$. On remarquera que tout chemin dans un tel automate est acceptant. Un automate *déterministe* sera ici définit comme un automate tel que $\forall q \in Q, \forall \alpha \in \Sigma, |\{q' \in Q \mid (q, \alpha, q') \in \delta\}| \leq 1$ mais toujours avec $Q = I = F$. On remarquera que dans le cadre de notre étude, un automate complet reconnait forcément Σ^*.

1.1 Espace sofique de Type Fini

On dit qu'un langage $L \subseteq \Sigma^*$ est de type fini s'il existe un ensemble fini $I \subseteq \Sigma^*$ de "facteurs interdits" tel que $L = \Sigma^* \setminus \Sigma^* I \Sigma^*$. Un automate $\mathcal{A} = \langle \Sigma, Q, \delta \rangle$ est dit (n, d)-*local* si pour toute paire de chemins dans $\mathcal{A}$:

$$q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} \dots \xrightarrow{a_n} q_n \qquad\qquad q_0' \xrightarrow{a_1} q_1' \xrightarrow{a_2} \dots \xrightarrow{a_n} q_n'$$

de longeur n et de même étiquettage, on a $q_d = q_d'$. $\mathcal{A}$ est dit local s'il est (n, d)-local pour des certains $n, d \in \mathbb{N}^*$. Un automate est dit *fortement connexe* si le graphe de ses états et transitions l'est.

(7) Soit $\mathcal{A}$ (n, d)-local, montrer que $\mathcal{A}$ est $(n+1, d)$-local et $(n+1, d+1)$-local.

(8) Montrer que tout langage L de type fini est reconnu par un automate local fortement connexe. Pour ce faire, pour $k = \max\{|u| : u \in I\}$, on pourra considérer un automate dont les états forment un sous-ensemble des mots de Σ^k.

(9) Montrer que si $\mathcal{A}$ est déterministe et fortement connexe, alors $\mathcal{A}$ est local ssi il n'existe pas de mot w et deux états distincts q, q' tel que $q \xrightarrow{w} q$ et $q' \xrightarrow{w} q'$. On pourra considerer un chemin de longeur $2|Q|^2$ dans l'automate produit.

(10) Proposer un algorithme polynomial pour tester la condition de la question précédente.

(11) Soit $\mathcal{A}$ un automate (n, d)-local fortement connexe.
Montrer que $L(\mathcal{A}) = \Sigma^* \setminus \Sigma^* I \Sigma^*$ pour

$$I := \left\{ w \in \Sigma^{n+1} \mid \neg\, \exists p, q \in Q, p \xrightarrow{w} q \right\}$$

Un espace X est dit de *type fini* si $F(X)$ est de type fini.

(12) Montrer que si $F(X)$ est reconnu par un automate fortement connexe alors il existe un automate déterministe reconnaissant $F(X)$. On veillera à ce que l'automate résultant soit local si et seulement si celui de départ l'est.

(13) Montrer qu'il existe un algorithme qui décide si un espace sofique donné par un automate fortement connexe est de type fini ou non.

1.2 Entropie d'un espace sofique

Pour X un espace de shift, on pose $F_n(X) = F(X) \cap \Sigma^n$, et on définit son entropie par

$$H(X) = \lim_{n \to \infty} \frac{\log_{|\Sigma|}(|F_n(X)|)}{n}$$

(14) Montrer que si $(u_n) \in \mathbb{R}_+^{\mathbb{N}}$ est tel que $\forall n, m \in \mathbb{N}^*, u_{n+m} \leq u_n + u_m$, alors $\lim_{n \to \infty} \frac{u_n}{n}$ existe.

(15) Montrer que $H(X)$ existe pour tout espace de shift X.

Sur $\Sigma = \{a, b\}$, on pose $\mathrm{G} = \left\{u \in \Sigma^{\mathbb{N}} \mid \forall i \in \mathbb{N}, u_i = a \Rightarrow u_{i+1} = b\right\}$

(16) Montrer que G est un espace de shift sofique de type fini.

(17) Calculer $H(\mathrm{G})$

* * *

2 Espaces Minimaux

Soit $u \in \Sigma^{\mathbb{N}}$. On définit $F(u) := F\left(\overline{\mathcal{O}(u)}\right)$ et $F_n(u) = F(u) \cap \Sigma^n$ et on pose $p_n(u) = |F_n(u)|$. Pour $w \in F(u)$, on pose $w^+ = \{v \in F(u) \mid \exists \alpha \in \Sigma, v = w\alpha\}$

(18) Montrer que s'il existe $n \in \mathbb{N}$ tel que $p_n(u) \leq n$ alors u est ultimement péridique. On pourra montrer que $p_{n+1}(u) - p_n(u) = \sum_{w \in F_n(u)} (|w^+| - 1)$.

(19) Montrer que u est ultimement périodique si et seulement s'il existe $C \in \mathbb{N}$ tel que $\forall n, p_n(u) \leq C$.

Un espace de shift X est dit minimal si pour tout $Y \subseteq X$ avec Y un espace de shift non vide, on a $Y = X$. Un $u \in \Sigma^{\mathbb{N}}$ est dit minimal si $\overline{\mathcal{O}(u)}$ est minimal.

(20) Montrer que X est minimal si et seulement si $\forall u, v \in X, \overline{\mathcal{O}(u)} = \overline{\mathcal{O}(v)}$

2.1 Morphismes et Substitution de Fibonacci

Un *morphisme* $f : \Sigma^* \longrightarrow \Sigma^*$ est une application telle que $f(\varepsilon) = \varepsilon$ et $\forall u, v \in \Sigma^*, f(uv) = f(u)f(v)$. On remarquera que f est uniquement définit à partir des valeurs de $f(\alpha)$ pour chaque $\alpha \in \Sigma$. Pour $u \in \Sigma^{\mathbb{N}}$ et f un morphisme, on définit $f(u) = f(u_0)f(u_1)f(u_2)...$

On dit qu'un morphisme f est une *substitution* si $|f(\alpha)| \geq 1$ pour tout $\alpha \in \Sigma$. On dit qu'une substition f est *α-prolongeable* si α est un préfixe strict de $f(\alpha)$. On dit que $u \in \Sigma^{\mathbb{N}}$ est un point-fixe de f si $f(u) = u$.

(21) Montrer que si f est une substition α-prolongeable, alors il existe un unique mot $u \in \Sigma^{\mathbb{N}}$ dont tous les $f^n(\alpha)$ sont préfixes. Montrer de plus qu'un tel u est un point fixe de f. On notera dans ce cas $f^\infty(\alpha)$ pour dénoter cet unique u.

(22) Montrer que $w \in F(f^\infty(\alpha)) \Leftrightarrow \exists l, r \in \Sigma^*, \exists k \in \mathbb{N}, lwr = f^k(\alpha)$

On se place dans le cas $\Sigma = \{a, b\}$. On définit *la substitution de Fibonacci* τ par $\tau(a) = ab$ et $\tau(b) = a$. On note $\text{fib} = \tau^\infty(a)$ qui existe par la question 21.

(23) Donner une expression de $\tau^{n+2}(a)$ en fonction de $\tau^{n+1}(a)$ et de $\tau^n(a)$. En déduire que $bb \notin F(\text{fib})$ et $aaa \notin F(\text{fib})$

(24) Soit $x \in \Sigma^*$, montrer par récurrence sur $|x|$ que $axa \notin F(\text{fib})$ ou $bxb \notin F(\text{fib})$

On définit $\overline{\tau}$ le *morphisme miroir* de τ par $\overline{\tau}(a) = ba$ et $\overline{\tau}(b) = a$.

(25) Montrer que si $u \in F(\text{fib})$, alors $a\overline{\tau}(u) = \tau(u)a$.

(26) Soit $g_2, g_3 := \varepsilon$ et $\forall n > 3, g_n := \tau^{n-3}(a)...\tau^2(a)\tau^1(a)a$. Montrer que

$$\forall n \geq 2, \tau^{n+2}(a) = \begin{cases} g_n\overline{\tau^n(a)}\ \overline{\tau^n(a)}ab \text{ si } n \text{ impair} \\ g_n\overline{\tau^n(a)}\ \overline{\tau^n(a)}ba \text{ si } n \text{ pair} \end{cases}$$

Pour $u \in \Sigma^\mathbb{N}$, on pose $D_n(u) = \{w \in \Sigma^n \mid \forall \alpha \in \Sigma, w\alpha \in F(u)\}$ et $G_n(u) = \{w \in \Sigma^n \mid \forall \alpha \in \Sigma, \alpha w \in F(u)\}$, respectivement les *facteurs spéciaux droit* et *gauche*. Remarquer que $D_n(u)$ est nécéssairement inclut dans les suffixes de $D_{n+1}(u)$.

(27) Montrer que $\forall n \in \mathbb{N}, |D_n(\text{fib})| = 1$. Pour ce faire, on pourra regarder la première lettre de $\overline{\tau^n(a)}$

(28) En déduire que $\forall n \in \mathbb{N}, p_n(\text{fib}) = n + 1$

(29) Montrer que si $u \in F(\text{fib})$, alors $\overline{u} \in F(\text{fib})$ avec $\overline{u}$ le miroir de u.

On verra dans la prochaine partie que beaucoup des propriétés vraies pour le mot infini de Fibonacci sont généralisables aux mots sturmiens.

2.2 Charactérisation des substitutions minimales

Une substitution f est dite *primitive* s'il existe $k \in \mathbb{N}$ tel que $\forall \alpha, \beta \in \Sigma, |f^k(\alpha)|_\beta > 0$. Un $u \in \Sigma^\mathbb{N}$ est dit *uniformément récurrent* si pour tout n, il existe N tel que chaque facteur de longeur N contient tous les facteurs de longeur n. Intuitivement, u est uniformément récurrent si tout facteur de longeur n apparait régulièrement à des intervalles bornés.

(30) Soit f une substitution primitive, montrer qu'il existe $k \in \mathbb{N}$ tel que f^k admette un point fixe.

(31) Montrer que si u est un point fixe d'une substitution primitive, alors u est uniformément récurrent.

(32) Montrer que u est uniformément récurrent si et seulement si $\overline{\mathcal{O}(u)}$ est minimal.

Comme fib est un point fixe d'une substitution primitive, on en déduit que $\overline{\mathcal{O}(\text{fib})}$ est minimal.

* * *

3 Mots sturmiens

Un mot $u \in \Sigma^{\mathbb{N}}$ est sturmien si $\forall n \in \mathbb{N}, p_n(u) = n + 1$. En particulier, $p_1(u) = 2$, on supposera donc que $\Sigma = \{a, b\}$ à partir de maintenant et ce jusqu'à la fin. La question 18 nous indique que ce sont les mots infinis "les plus proches d'être périodique".

(33) Montrer que si u est sturmien, alors $|D_n(u)| = 1$. On notera alors $d_n(u)$ l'unique mot de $D_n(u)$

Un ensemble $A \subseteq \Sigma^*$ de mots est dit *équilibré* s'il est clot par facteur (pour tout $u \in A$ et v facteur de u, on a $v \in A$), et que, pour tout $v, w \in A$ de même longeur, pour tout $\alpha \in \Sigma$,

$$| \ |v|_\alpha - |u|_\alpha \ | \leq 1$$

Soit $u \in \Sigma^{\mathbb{N}}$.

(34) Montrer que $F(u)$ n'est pas équilibré si et seulement s'il existe $v \in \Sigma^*$ tel que $ava \in F(u)$ et $bvb \in F(u)$.

(35) Montrer que la question précédente reste vraie avec v un palindrome. Précisément, montrer que $F(u)$ n'est pas équilibré si et seulement s'il existe un palindrome $v \in \Sigma^*$ tel que $ava, bvb \in F(u)$.

(36) Montrer que si A est équilibré, alors $\forall n \in \mathbb{N}, |A \cap \Sigma^n| \leq n + 1$.

(37) Soient $w \in F(u)$ et $n \in \mathbb{N}$ avec u sturmien. Montrer que si $|w| > 2n$, alors $d_n(u)$ est un facteur de w.

(38) Montrer que $F(u)$ est équilibré et u est non ultimement périodique si et seulement si u est sturmien.

(39) Soit u un mot sturmien, montrer que $F(u)$ est clot par miroir, c'est à dire que si $w \in F(u)$ alors $\overline{w} \in F(u)$ pour $\overline{w_1 ... w_n} = w_n ... w_1$

* * *

4 Théorème de Mignosi et Séébold

On dit qu'un morphisme f est un morphisme *sturmien* si pour tout mot sturmien $u \in \Sigma^{\mathbb{N}}$, $f(u)$ est aussi un mot sturmien.

(40) Montrer que la composition de 2 morphismes sturmiens est un morphisme sturmien.

Pour rappel, on a noté τ le morphisme de fibonacci (définit avant la question 26) et $\overline{\tau}$ le morphisme miroir de fibonacci (définit avant la question 30). On définit aussi le morphisme E tel que $E(a) = b$ et $E(b) = a$ et Id le morphisme identité définit par $\mathrm{Id}(a) = a$ et $\mathrm{Id}(b) = b$.

(41) Montrer que Id, E, τ sont des morphismes sturmiens.

(42) Montrer que si $\tau(u)$ est sturmien, alors u l'est.

On admettera que les deux questions précédentes sont vraies pour $\overline{\tau}$ aussi, les preuves étant très similaire à celle de τ.

Un morphisme sturmien f est dit a-sturmien si $f(ab)$ ou $f(ba)$ contient aa en facteur. Il est dit b-sturmien si $f(ab)$ ou $f(ba)$ contient bb en facteur.

(43) Montrer que si $f \notin \{E, \mathrm{Id}\}$ est un morphisme sturmien, alors il est a-sturmien ou b-sturmien.

(44) Soit f un morphisme sturmien autre que Id, E, alors il existe $g \in \{\tau, \overline{\tau}, E \circ \tau, E \circ \overline{\tau}\}$ et h sturmien tel que $f = g \circ h$

On définit la taille d'un morphisme sturmien f par $\|f\| := |f(a)| + |f(b)|$.

(45) Montrer qu'un morphisme f est sturmien si et seulement s'il est une suite de compositions de Id, E, τ et $\overline{\tau}$.

* * *

Fin du sujet.

* * *

Responsables

Sujet écrit en typst par **Coda Bourotte** (@cypooos, https://bourotte.com), étudiante à l'ENS de Lyon. Mise en place du site web et trailer par elle aussi.

Sujet testé par:

- **Sarah Monestier**, ENS Lyon (@sarenard)
- **Elowan H**, ENS Lyon (https://www.elowarp.fr/)

Merci à Florent Ferrari pour la correction des nombreuses fautes d'orthographe.

Nous ne sommes pas affiliées au concours des ENS de quelconque manière que ce soit.