Epreuve d' informatique
durée 4h

Concours 2001

Les correcteurs attendent des réponses précises et concises aux questions posées. On demande à plusieurs reprises de proposer des algorithmes. On exprimera ces algorithmes avec un point de vue de haut niveau, sans décrire leur implantation effective (cf l'algorithme proposé dans l'énoncé de la question 2.9). La complexité d'un algorithme doit toujours être interprétée comme le nombre d'opérations élémentaires (calculs, comparaisons, ...) nécessaires à son exécution.

La partie 4 est largement indépendante des parties précédentes. En règle générale, les références à un résultat d'une autre partie sont explicitement mentionnées.

Soit A un alphabet, c'est à dire un ensemble fini non vide. On note A^* l'ensemble des mots finis formés de lettres de A, y compris le mot vide ε, et $A^+ = A^* \setminus \{\varepsilon\}$. Si u et v sont deux mots, on note uv le mot obtenu par concaténation de u et v ; l'ensemble A^* muni de cette loi de composition est un monoïde. On note $|w|$ la longueur d'un mot w.

On dit qu'un mot v est *facteur* d'un autre mot w s'il existe des mots x et y tels que $w = xvy$. Si de plus on peut prendre $x = \varepsilon$, le mot v est dit *préfixe* de w, tandis que si $y = \varepsilon$, le mot v est dit *suffixe* de w.

Dans les exemples, on utilisera le plus souvent les alphabets $A_1 = \{a\}$, $A_2 = \{a, b\}$ et $A_3 = \{a, b, c\}$.

L'ensemble des lettres de A qui apparaissent effectivement dans un mot $w \in A^*$ est noté $\mathrm{Alph}(w)$. Le cardinal d'un ensemble S est noté $\mathrm{Card}(S)$. On note $S \subset T$ quand l'ensemble S est inclus (au sens large) dans l'ensemble T.

1 Morphismes et L-systèmes

Soient A et B deux alphabets. Un *morphisme* de A^* dans B^* est une application $f\colon A^* \to B^*$ telle que, pour tous mots u et v dans A^*, $f(uv) = f(u)f(v)$.

Question 1.1.

Montrer qu'un morphisme $f\colon A^* \to B^*$ est entièrement défini par la donnée de $f(x)$ pour chaque lettre $x \in A$.

Cette observation permet d'exprimer les morphismes de manière plus compacte. Ainsi, on notera $f = (a \mapsto aab, b \mapsto ba, c \mapsto \varepsilon)$ l'unique morphisme de A_3^* dans A_2^* tel que $f(a) = aab$, $f(b) = ba$ et $f(c) = \varepsilon$.

Un morphisme $f\colon A^* \to B^*$ est dit *non-effaçant* si $f(x) \neq \varepsilon$ pour tout $x \in A$, *effaçant* dans le cas contraire ; il est dit *lettre-à-lettre* si $f(x) \in B$ pour tout $x \in A$.

Si f est un morphisme de A^* dans lui-même, on note $f^1 = f$, $f^2 = f \circ f$ et plus généralement $f^{n+1} = f^n \circ f$.

On appelle *D0L-système* un triplet $G = (A, f, u_0)$, où A est un alphabet, f est un morphisme de A^* dans A^* et $u_0 \in A^*$ (u_0 est appelé l'*axiome* du D0L-système G).

À chaque D0L-système $G = (A, f, u_0)$, on associe la suite infinie de mots $S(G) = (u_0, u_1, u_2, \ldots)$ telle que u_0 est l'axiome de G et $u_{n+1} = f(u_n)$ pour tout $n \in \mathbb{N}$, ce qu'on peut noter $u_n = f^n(u_0)$. On associe aussi à G le langage $L(G) = \{u_0, u_1, u_2, \ldots\} = \{f^n(u_0) : n \in \mathbb{N}\}$.

Par exemple, soit $G_1 = (A_2, (a \mapsto b, b \mapsto aa), a)$. On a

$$S(G_1) = (a, b, aa, bb, aaaa, bbbb, \ldots)$$

et $L(G_1) = \{a^{2^i}, b^{2^i} : i \in \mathbb{N}\}$.

Question 1.2.

Construire un D0L-système G_2 tel que $L(G_2) = L(G_1)$ mais $S(G_2) \neq S(G_1)$.

Question 1.3.

Soit θ le morphisme de A_2^* dans A_2^* défini par $\theta(a) = ab$ et $\theta(b) = ba$. On considère le D0L-système $T = (A_2, \theta, a)$. Calculer les cinq premiers termes de $S(T)$. En utilisant un morphisme lettre-à-lettre $\iota\colon A_2^* \to A_2^*$, donner une formule simple permettant de passer de $\theta^n(a)$ à $\theta^{n+1}(a)$. Quelle est la longueur de $\theta^n(a)$? En déduire que $L(T)$ n'est pas un langage rationnel.

On appelle *HD0L-système* un quintuplet $H = (B, f, u_0, A, g)$, où A et B sont des alphabets (éventuellements égaux), (B, f, u_0) est un D0L-système que l'on notera H^0, et g est un morphisme de B^* dans A^*. Si $S(H^0) = (u_0, u_1, u_2, \ldots)$ est la suite de mots associée à H^0, alors on associe à H la suite $S(H) = (g(u_0), g(u_1), g(u_2), \ldots)$ et le langage $L(H) = \{g(f^n(u_0)) : n \in \mathbb{N}\}$.

Question 1.4.

Soit $H = (A_2, (a \mapsto ab, b \mapsto b), a, A_1, (a \mapsto a, b \mapsto a))$. Montrer que $L(H) = A_1^+$. Existe-t-il un D0L-système G tel que $L(G) = L(H)$?

Note historique : les D0L-systèmes et les HD0L-systèmes, ainsi que d'autres systèmes similaires permettant de construire des langages, sont collectivement appelés L-systèmes. Ils ont été introduits en 1968 par Aristid Lindenmayer pour modéliser la croissance de certains organismes vivants.

2 Mots infinis engendrés par L-systèmes

Soit A un alphabet. Un *mot infini* sur A est une suite $\mathbf{y} = (y_i)_{i\in\mathbb{N}}$ à valeurs dans A. On note $A^{\mathbb{N}}$ l'ensemble de tous les mots infinis sur A. Un *préfixe* de $\mathbf{y}$ est un mot (fini) de la forme $y_0y_1\dots y_{k-1}$, avec $k \in \mathbb{N}$, et un *facteur* de $\mathbf{y}$ est un mot de la forme $y_iy_{i+1}\dots y_{i+k-1}$, avec $i, k \in \mathbb{N}$.

Soit $L \subset A^*$ un langage et $\mathbf{y} \in A^{\mathbb{N}}$ un mot infini. On dit que L *engendre* le mot infini $\mathbf{y}$ si les deux conditions suivantes sont vérifiées :

(i) L est infini ;

(ii) tout élément de L est préfixe de $\mathbf{y}$.

Question 2.1.

Montrer que si le langage L engendre deux mots infinis $\mathbf{y}$ et $\mathbf{z}$, alors $\mathbf{y} = \mathbf{z}$. Montrer que quel que soit le mot infini $\mathbf{y}$, il existe au moins un langage qui engendre $\mathbf{y}$.

Question 2.2.

Montrer qu'un langage L engendre un mot infini si et seulement si L est infini et pour tout couple (u, v) d'éléments de L, soit u est préfixe de v, soit v est préfixe de u.

Si G est un D0L-système (ou un HD0L-système), et si $L(G)$ engendre un mot infini $\mathbf{y}$, on dit aussi que G engendre $\mathbf{y}$, et on note $\mathbf{y} = W(G)$.

Dans les questions 2.3, 2.5, 2.6 et 2.7, il est demandé de construire un D0L-système ayant certaines propriétés ; à chaque fois, un seul exemple suffit : on ne cherchera pas à caractériser tous les D0L-systèmes ayant les propriétés requises ni à prouver l'unicité de l'exemple construit.

Question 2.3.

Donner un D0L-système Q engendrant le mot infini $\mathbf{q} = (q_i)_{i\in\mathbb{N}} \in A_2^{\mathbb{N}}$ défini par $q_{2i} = a$ et $q_{2i+1} = b$ pour tout $i \in \mathbb{N}$.

Question 2.4.

Soit $G = (A, f, u_0)$ un D0L-système tel que le langage associé $L(G)$ est infini. Montrer que G engendre un mot infini si et seulement si u_0 est préfixe de $f(u_0)$.

Question 2.5.

Donner un D0L-système K_1 tel que $S(K_1) = (u_0, u_1, u_2, \ldots)$ avec $u_0 \neq \varepsilon$, $u_1 \neq \varepsilon$, $u_2 \neq \varepsilon$ et $u_3 = \varepsilon$. Que vaut $L(K_1)$? Est-ce que K_1 engendre un mot infini ?

Question 2.6.

Donner un D0L-système K_2 tel que $S(K_2) = (u_0, u_1, u_2, \ldots)$ avec u_0 préfixe de u_1, $u_0 \neq u_1$, $u_1 \neq u_2$ mais $u_2 = u_3$. Que vaut $L(K_2)$? Est-ce que K_2 engendre un mot infini ?

Question 2.7.

Donner un D0L-système K_3 tel que $L(K_3)$ est infini mais n'engendre pas de mot infini.

Soit $f\colon A^* \to A^*$ un morphisme, et $x \in A$ une lettre. On dit que x est une lettre *mortelle* (pour f) s'il existe un entier $n \geq 1$ tel que $f^n(x) = \varepsilon$, et que x est une lettre *immortelle* dans le cas contraire.

Question 2.8.

Quelles sont les lettres immortelles dans l'exemple K_2 de la question 2.6 ? Montrer que si $G = (A, f, u_0)$ est un D0L-système tel que $f(u_0) = u_0 v$ avec $v \in A^*$, alors $L(G)$ est infini si et seulement si le mot v contient une lettre immortelle pour f.

Question 2.9.

L'algorithme suivant prend en entrée un alphabet A et un morphisme $f\colon A^* \to A^*$, et retourne l'ensemble des lettres mortelles pour f. Si $w \in A^*$ est un mot, on note $w[i]$ la lettre de rang i de w, de sorte que $w = w[0]w[1]\dots w[|w|-1]$.

Lettres-mortelles (A, f)
    $T \leftarrow \emptyset$
    $M \leftarrow \emptyset$
    pour tout $x \in A$ **faire**
        pour tout $y \in A$ **faire**
            $N[x,y] \leftarrow 0$
    pour tout $y \in A$ **faire**
        $w \leftarrow f(y)$
        pour i **de** 0 **à** $|w|-1$ **faire**
            $N[w[i],y] \leftarrow N[w[i],y] + 1$
        $L[y] \leftarrow |w|$
        si $L[y] = 0$
            alors $T \leftarrow T \cup \{y\}$
    tant que $T \neq \emptyset$ **faire**
        choisir $x \in T$
        $T \leftarrow T \setminus \{x\}$
        $M \leftarrow M \cup \{x\}$
        pour tout $y \in A \setminus (M \cup T)$ **faire**
            $L[y] \leftarrow L[y] - N[x,y]$
            si $L[y] = 0$
                alors $T \leftarrow T \cup \{y\}$
    retourner M

Justifier la validité de cet algorithme (on précisera notamment la signification de la matrice N). Montrer que sa complexité est $O(k^2 + m)$, où $k = \mathrm{Card}(A)$ et $m = \sum_{x \in A} |f(x)|$.

Question 2.10.

Proposer et justifier un algorithme prenant en entrée un D0L-système G et un entier ℓ, retournant le préfixe de longueur ℓ de $W(G)$ si G engendre un mot infini, et retournant ń G n'engendre pas de mot infini ż sinon.

Soit $f\colon A^* \to B^*$ un morphisme non-effaçant. Étant donné un mot infini $\mathbf{y} \in A^{\mathbb{N}}$, le langage $\{f(w)\colon w \text{ préfixe de } \mathbf{y}\}$ engendre un unique mot infini, que l'on notera $f(\mathbf{y})$. On prolonge ainsi f en une application de $A^{\mathbb{N}}$ dans $B^{\mathbb{N}}$.

Soit $f\colon A^* \to A^*$ un morphisme non-effaçant et $\mathbf{y} \in A^{\mathbb{N}}$ un mot infini. On dit que $\mathbf{y}$ est un *point fixe non trivial* de f si $\mathbf{y} = f(\mathbf{y})$ et s'il existe une lettre $x \in A$ qui apparaît dans $\mathbf{y}$ et telle que $f(x) \neq x$.

Question 2.11.

Soit $\zeta = (a \mapsto ab, b \mapsto ab)$. *Donner un point fixe non trivial de* ζ. *Le morphisme* ζ *a-t-il un autre point fixe non trivial ?*

Question 2.12.

Soit $\eta = (a \mapsto aba, b \mapsto b)$. *Donner deux points fixes non triviaux de* η. *Le morphisme* η *a-t-il d'autres points fixes non triviaux ?*

Question 2.13.

Montrer que si $\mathbf{y}$ *est point fixe non trivial d'un morphisme non-effaçant* $f \colon A^* \to A^*$, *alors* $\mathbf{y}$ *est engendré par un D0L-système que l'on précisera. En déduire que si* $f(x) \neq x$ *pour tout* $x \in A$, f *a au plus* $\mathrm{Card}(A)$ *points fixes non triviaux.*

On dit qu'un mot infini $\mathbf{y} = (y_i)_{i \in \mathbb{N}}$ est *ultimement périodique* s'il existe des entiers $i_0 \geq 0$ et $p \geq 1$ tels que pour tout $i \geq i_0$, $y_i = y_{i+p}$.

Question 2.14.

Parmi les exemples de D0L-systèmes déjà construits, en donner un qui engendre un mot infini ultimement périodique. Montrer que tout mot infini ultimement périodique peut être engendré par un HD0L-système.

Question 2.15.

Soit $T = (A_2, \theta, a)$ *le D0L-système défini à la question 1.3. Montrer que* T *engendre un mot infini* $\mathbf{t} = W(T) = (t_i)_{i \in \mathbb{N}}$. *Comment calculer* t_i *en fonction de* i, *sans calculer tous les termes précédents comme le fait l'algorithme de la question 2.10 ? Montrer que* $\mathbf{t}$ *n'est pas ultimement périodique.*

Le mot infini $\mathbf{t}$ est appelé *mot infini de Thue-Morse.*

Question 2.16.

Soient $\mu = (a \mapsto abc, b \mapsto ac, c \mapsto b)$ *et* $\psi = (a \mapsto abb, b \mapsto ab, c \mapsto a)$. *Montrer que le HD0L-système* $T' = (A_3, \mu, a, A_2, \psi)$ *engendre aussi le mot infini de Thue-Morse.*

Question 2.17.

> *Soit $G = (A, f, u_0)$ un D0L-système. Pour n entier strictement positif, on note $G_n = (A, f^n, u_0)$. Montrer que si G_m et G_n engendrent des mots infinis, alors $W(G_m) = W(G_n)$.*

3 Hiérarchie

Soit $G = (A, f, u_0)$ un D0L-système. Si f est non-effaçant, on dit que G est un PD0L-système.

Soit $H = (B, f, u_0, A, g)$ un HD0L-système. Si g est non-effaçant, on dit que H est un ND0L-système. Si g est lettre-à-lettre, on dit que H est un CD0L-système.

On peut également combiner ces deux notations. Si f est non-effaçant, on dit que H est un HPD0L-système. Si f et g sont non-effaçants, on dit que H est un NPD0L-système. Si f est non-effaçant et g lettre-à-lettre, on dit que H est un CPD0L-système.

On a ainsi défini huit types de L-systèmes. Pour chaque type X, on note $\mathcal{W}_A(X)$ l'ensemble des mots infinis sur A engendrés par un X-système. Les huit classes de mots infinis ainsi définies vérifient de manière évidente les inclusions suivantes :

$$\begin{array}{ccccccc} \mathcal{W}_A(\text{D0L}) & \subset & \mathcal{W}_A(\text{CD0L}) & \subset & \mathcal{W}_A(\text{ND0L}) & \subset & \mathcal{W}_A(\text{HD0L}) \\ \cup & & \cup & & \cup & & \cup \\ \mathcal{W}_A(\text{PD0L}) & \subset & \mathcal{W}_A(\text{CPD0L}) & \subset & \mathcal{W}_A(\text{NPD0L}) & \subset & \mathcal{W}_A(\text{HPD0L}) \end{array} \tag{1}$$

Le but de cette partie est de voir lesquelles de ces inclusions sont strictes.

Dans les questions 3.1 et 3.2, on utilisera la propriété suivante, qui sera démontrée dans la partie 4 :

Proposition 1. *Le mot infini de Thue-Morse $\mathbf{t} = W(A_2, (a \mapsto ab, b \mapsto ba), a)$ (voir les questions 1.3 et 2.15) ne contient aucun facteur de la forme vvv, avec $v \in A_2^+$.*

Question 3.1.

> *Soit le D0L-système $J_1 = (A_3, (a \mapsto abccc, b \mapsto baccc, c \mapsto \varepsilon), a)$. Montrer que $W(J_1) \notin \mathcal{W}_{A_3}(PD0L)$. Est-il possible de construire un tel contre-exemple sur l'alphabet A_2 ?*
>
> Indication : observer d'abord qu'en effaçant les c dans $W(J_1)$, on retrouve le mot infini de Thue-Morse $\mathbf{t}$, puis que si $W(J_1)$ était engendré par un PD0L-système (A_3, f, u_0), on aurait nécessairement $f(c) = c$.

Question 3.2.

> *Soit le NPD0L-système $J_2 = (A_2, \theta, a, A_2, \varphi)$, où $\theta = (a \mapsto ab, b \mapsto ba)$ et $\varphi = (a \mapsto aa, b \mapsto bb)$. Montrer que $\varphi(\mathbf{t}) = W(J_2) \notin \mathcal{W}_{A_2}(D0L)$.*
>
> Indication : montrer que si $\varphi(\mathbf{t})$ était engendré par un D0L-système (A_2, f, u_0), alors il contiendrait un mot de la forme $vvvv$ avec $|v| \geq 2$.

Question 3.3.

Montrer que $\mathcal{W}_A(D0L) \subset \mathcal{W}_A(NPD0L)$. *En déduire que* $\mathcal{W}_A(ND0L) = \mathcal{W}_A(NPD0L)$ *et* $\mathcal{W}_A(HD0L) = \mathcal{W}_A(HPD0L)$.

Indication : on pourra procéder par récurrence sur la taille de l'alphabet, et montrer que si $\mathbf{y} = W(A, f, u_0)$ avec f effaçant, alors il existe un alphabet B de cardinal strictement inférieur à celui de A et des morphismes $g\colon A^* \to B^*$ et $h\colon B^* \to A^*$ tels que $\mathbf{y} = W(B, g \circ h, g(u_0), A, h)$.

Question 3.4.

Montrer que $\mathcal{W}_A(HPD0L) \subset \mathcal{W}_A(NPD0L)$.

Indication : on pourra commencer par montrer que, pour tout morphisme $f\colon B^* \to B^*$, il existe un entier strictement positif N tel que pour tout $x \in B$, $\mathrm{Alph}(f^N(x)) = \mathrm{Alph}(f^{2N}(x))$, puis utiliser la question 2.17 pour se ramener à un HPD0L-système $\tilde{H} = (B, \tilde{f}, u_0, A, \tilde{g})$ tel que $\tilde{g}(\tilde{f}(x)) = \varepsilon$ si et seulement si $\tilde{g}(x) = \varepsilon$.

Question 3.5.

Montrer que $\mathcal{W}_A(NPD0L) \subset \mathcal{W}_A(CPD0L)$. *Illustrer cette inclusion en construisant un CPD0L-système engendrant le mot infini* $\varphi(\mathbf{t})$ *construit à la question 3.2.*

Indication : si $H = (B, f, u_0, A, g)$ est le NPD0L-système de départ, on pourra, après avoir modifié f et g, utiliser l'alphabet intermédiaire $\tilde{B} = \{(x, i)\colon x \in B, 1 \leq i \leq |g(x)|\}$.

Question 3.6.

En rassemblant les résultats de cette partie, conclure en précisant la nature (inclusion stricte ou égalité) de toutes les inclusions figurant dans le diagramme (1). On distinguera les cas où $\mathrm{Card}(A)$ *vaut 1, 2, ou au moins 3.*

4 Mots sans carré, mots sans cube

Soit w un mot fini ou infini. On dit que w *contient un carré* s'il existe un mot non vide $v \in A^+$ tel que vv est facteur de w (le mot vv est appelé le *carré* de v). Dans le cas contraire, on dit que w est *sans carré*. Ainsi, *abcbacbab* contient un carré (le carré de *cba*) tandis que *abcacbabc* est sans carré. On note $E^2(A)$ l'ensemble des mots de A^* sans carré.

De même, on dit que w est *sans cube* s'il ne contient aucun facteur de la forme vvv avec $v \in A^+$.

Question 4.1.

Montrer qu'il n'existe qu'un nombre fini de mots sans carré dans A_2^*. *Décrire le langage* $E^2(A_2)$.

Question 4.2.

Proposer et justifier un algorithme prenant en entrée un mot w *et retournant ń* w *contient un carré ż ou ń* w *est sans carré ż selon la nature de* w. *Quelle est sa complexité ?*

Question 4.3.

Proposer et justifier un algorithme prenant en entrée l'alphabet A *et un entier* ℓ, *et retournant la liste des mots sans carré de longueur inférieure ou égale à* ℓ *dans* A^*. *La complexité devra être au plus en* $O(k\ell^2 m)$, *où* $k = \mathrm{Card}(A)$ *et* m *est le nombre de mots sans carré retournés.*

On pourra utiliser le fait qu'un mot w est sans carré si et seulement si w n'a aucun suffixe de la forme vv et son préfixe de longueur $|w| - 1$ est sans carré.

Dans les questions qui suivent, on cherche à montrer que $E^2(A_3)$ est infini. On considère pour cela le morphisme $\mu = (a \mapsto abc, b \mapsto ac, c \mapsto b)$.

Question 4.4.

Montrer que μ *est injectif.*

On note V l'ensemble des mots de A_3^* qui ne contiennent ni aa, ni bb, ni cc, ni aba, ni cbc comme facteurs.

Question 4.5.

Montrer que pour tout mot $w \in V$, $\mu(w) \in V$.

Question 4.6.

Montrer que pour tout mot $w \in A_3^*$ *et tout facteur* v *de* $\mu(w)$ *autre que* ε *ou* b, *il existe un unique triplet* (x, y, z), *où* $x \in \{\varepsilon, c, bc\}$, $y \in A_3^*$ *et* $z \in \{\varepsilon, a, ab\}$ *tel que* $v = x\mu(y)z$. *Montrer que le mot* y *est alors facteur de* w.

Question 4.7.

Montrer que, si $w \in A_3^*$ et $\mu(w)$ contient un carré, alors w contient soit un carré, soit un mot de la forme $aybya$ avec $y \in A_3^*$.

Indication : si $\mu(w)$ contient vv, commencer par appliquer le résultat de la question 4.6 à v.

Question 4.8.

Montrer qu'aucun mot de V ne contient de facteur de la forme $aybya$.

Question 4.9.

Déduire de ce qui précède que $\mu(V \cap E^2(A_3)) \subset V \cap E^2(A_3)$. Construire un D0L-système G tel que $L(G)$ est infini et $L(G) \subset E^2(A_3)$.

Un mot sans carré $w \in E^2(A_3)$ est dit *indéfiniment prolongeable* si pour tout ℓ, il existe u et v dans A_3^* de longueur ℓ tels que uwv soit sans carré. Un mot sans carré est dit *non prolongeable* si pour toute lettre $x \in A_3$, xw et wx contiennent chacun un carré.

Question 4.10.

Montrer qu'il existe dans A_3^* une infinité de mots sans carré indéfiniment prolongeables, et une infinité de mots sans carré non prolongeables (on commencera par contruire un mot sans carré non prolongeable).

Question 4.11.

En utilisant le résultat de la question 2.16, montrer que le mot infini de Thue-Morse $\mathbf{t}$ *est sans cube.*