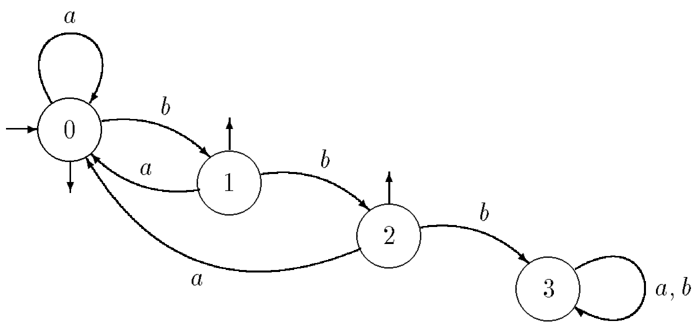
Option Informatique en Spé MP et MP*
Extractions conservant la rationalité : le corrigé

L'effaceur d'indices pairs

Question 1 • Un mot u appartient à T ssi il peut s'écrire $v_1 a v_2 a \dots v_n a v_{n+1}$, chaque v_i étant égal à ε, b ou bb. Donc T est décrit par l'expression rationnelle $(a+ba+bba)^*(\varepsilon+b+bb)$.

Question 2 • La lecture de trois b consécutifs amène dans l'état poubelle. Tout autre état de cet automate déterministe complet est final.



Question 3 • $\varphi(T) = \Sigma^*$, et est donc rationnel. Preuve : $\varepsilon = \varphi(\varepsilon)$, or $\varepsilon \in L$; donc $\varepsilon \in \varphi(L)$. Soit maintenant $u = u_1 u_2 \dots u_n \in \Sigma^*$, avec $n \geqslant 1$. Le mot $v = u_1 a u_2 a \dots u_n a$ appartient à T (on ne risque pas d'y trouver le facteur bbb, puisqu'il ne contient même pas le facteur bb) et $\varphi(v) = u$: donc $\Sigma^* \subset \varphi(T)$. L'inclusion inverse est banale.

Question 4 • Considérons l'automate $\mathcal{A}' = (Q', \delta', i', F')$ défini comme suit :

- $Q' = Q \times \{0,1\}$;
- $i' = (i, 0)$;
- $F' = F \times \{0,1\}$;
- $\delta'\big((q,0),x\big)$ est construit en ajoutant l'état $\big(\delta(q,x),1\big)$ à l'ensemble des états $(q'',0)$ pour lesquels il existe $q' \in Q$ et $y \in \Sigma$ tels que $\delta(q,x) = q'$ et $\delta(q',y) = q''$.

Remarquons que l'automate $\mathcal{A}'$ n'est pas nécessairement déterministe. Prouvons que $\mathcal{A}'$ reconnaît $\varphi(L)$.
• Montrons d'abord que tout mot u reconnu par $\mathcal{A}'$ appartient à $\varphi(L)$. Si $u = \varepsilon$, alors $i' = (i,0) \in F'$, donc $i \in F$, si bien que ε est reconnu par $\mathcal{A}$; alors $\varepsilon \in L$, et finalement $\varepsilon \in \varphi(L)$. Soit maintenant $u = u_1 u_2 \dots u_n$ reconnu par $\mathcal{A}'$, avec $n \geqslant 1$. Observons un calcul réussi de $\mathcal{A}'$, d'étiquette u :

$$s_0 \xrightarrow{u_1} s_1 \dots \xrightarrow{u_n} s_n$$

Distinguons deux cas de figure selon que la deuxième composante de s_n est égale à 0 ou à 1. Dans le premier cas, il existe une suite $(q_i)_{0 \leqslant i \leqslant 2n}$ d'états, et un mot $v = v_1 v_2 \dots v_n$ tels que :

- $q_0 = i$;
- $q_{2i-2} \xrightarrow{u_i} q_{2i-1}$, $q_{2i-1} \xrightarrow{v_i} q_{2i}$ et $(q_{2i}, 0) = s_i$ pour $0 \leqslant i \leqslant n$;
- $q_{2n} \in F$.

Ceci montre que le mot $w = u_1 v_1 u_2 v_2 \dots u_n v_n$ est l'étiquette d'un calcul réussi de $\mathcal{A}$; donc w appartient à L, et par suite $u = \varphi(w)$ appartient à $\varphi(L)$. Dans le deuxième cas, il existe une suite $(q_i)_{0 \leqslant i < 2n}$ d'états, et un mot $v = v_1 v_2 \dots v_{n-1}$ tels que :

- $q_0 = i$;
- $q_{2i-2} \xrightarrow{u_i} q_{2i-1}$ pour $0 \leqslant i \leqslant n$;
- $q_{2i-1} \xrightarrow{v_i} q_{2i}$ pour $0 \leqslant i < n$;

- $(q_{2i}, 0) = s_i$ pour $0 \leqslant i < n$ et $(q_{2n-1}, 1) = s_n$;
- $q_{2n} \in F$.

On met cette fois en évidence un mot $v = v_1 v_2 \dots v_{n-1}$ tel que $w = u_1 v_1 u_2 v_2 \dots u_n$ soit un calcul réussi de $\mathcal{A}$; comme dans le premier cas, $w \in L$, donc $u = \varphi(w)$ appartient à $\varphi(L)$.

• Montrons maintenant que tout mot u de $\varphi(L)$ est reconnu par $\mathcal{A}'$. Si $\varepsilon \in \varphi(L)$, alors $\varepsilon \in L$ puisque c'est le seul mot dont l'image par φ est ε ; alors ε est reconnu par $\mathcal{A}$, donc $i \in F$, d'où $i' = (i, 0) \in F'$ ce qui prouve que ε est reconnu par $\mathcal{A}'$. Considérons maintenant un mot $u \neq \varepsilon$ appartenant à $\varphi(L)$; notons $n = |u|$, $u = u_1 u_2 \dots u_n$. Soit w un mot de L tel que $u = \varphi(w)$. Nous distinguons deux cas selon que la longueur de w est paire ou impaire. Dans le premier cas, $w = u_1 v_1 \dots\dots u_n v_n$ est l'étiquette d'un calcul réussi de $\mathcal{A}$:

$$q_0 \xrightarrow{u_1} q_1 \xrightarrow{v_1} q_2 \dots \xrightarrow{u_n} q_{2n-1} \xrightarrow{v_n} q_{2n}$$

Nous noterons alors $s_i = (q_{2i}, 0)$ pour $0 \leqslant i \leqslant n$. Dans le deuxième cas, $w = u_1 v_1 \dots\dots u_n$ est l'étiquette d'un calcul réussi de $\mathcal{A}$:

$$q_0 \xrightarrow{u_1} q_1 \xrightarrow{v_1} q_2 \dots \xrightarrow{u_n} q_{2n-1}$$

Nous noterons cette fois $s_i = (q_{2i}, 0)$ pour $0 \leqslant i < n$ et $s_n = (q_{2n}, 1)$. Nous constatons que, dans les deux cas, u est l'étiquette d'un calcul réussi de $\mathcal{A}'$, à savoir : $s_0 \xrightarrow{u_1} s_1 \xrightarrow{u_2} s_1 \dots \xrightarrow{u_n} s_n$.

Question 5 • Supposons L rationnel. Le lemme de l'étoile affirme l'existence d'une constante N telle que tout mot u de L, de longueur au moins égale à N, se décompose en $u = rst$ avec $|rs| \leqslant N$, $s \neq \varepsilon$ et $rs^k t \in L$ pour tout $k \in \mathbb{N}$; en particulier, le mot rt devrait appartenir à L. Soit $u = v_{4N^2}$. Le mot u appartient à L, et $|u| = 4N^2 \geqslant N$; on peut donc écrire $u = rst$ où $r = x(1) \dots x(i)$, $s = x(i+1) \dots x(j)$ et $t = x(j+1) \dots x(4N^2)$, avec $i + 1 \leqslant j \leqslant N$. Observons le mot $rt = x(1) \dots x(i) x(j+1) \dots x(4N^2)$: $|rt| = |u| - |s| = 4N^2 - (j - i)$; $0 < j - i \leqslant N$, donc $4N^2 - N \leqslant |rt| < 4N^2 = (2N)^2$; mais $4N^2 - N \geqslant 4N^2 - 4N + 1 = (2N-1)^2$. Donc $|rt|$ n'est pas un carré parfait ; à plus forte raison, ce n'est pas un carré parfait pair. Or $|t| = |u| - |rs| \geqslant 4N^2 - N > 0$, donc la dernière lettre de rt est celle de t, soit celle de rst, qui est $x(4N^2) = a$ puisque $4N^2$ est un carré parfait pair. On en déduit que rt ne peut appartenir à L, d'où la contradiction.

Question 6 • Dans un mot de L, les a n'apparaissent qu'à des positions paires ; ils sont donc tous effacés par φ ; de plus, $\varepsilon \notin \varphi(L)$ puisque tout mot de ce langage commence par un a. Ceci montre que $\varphi(L)$ est contenu dans $\{b\}^+$. Réciproquement, si $n \geqslant 1$, alors b^n est l'image par φ du mot $x(2n)$. Conclusion : $\varphi(L) = \{b\}^+$.

Parties ultimement périodiques de $\mathbb{N}^*$

Question 7 • Stabilité par complémentation : si $n \in S \iff n + p \in S$, alors $n \notin S \iff n + p \notin S$. Donc $\mathbb{N}^* \setminus S$ est une p.u.p. de $\mathbb{N}^*$.

• Soient S et T deux p.u.p. ; soient $n_0 \geqslant 1$ et $p \geqslant 1$ tels que, pour $n \geqslant n_0$, $n \in S \iff n + p \in S$. Soient de même $n_1 \geqslant 1$ et $q \geqslant 1$ tels que, pour $n \geqslant n_1$, $n \in T \iff n + q \in T$. Alors, pour $n \geqslant \max(n_0, n_1)$:

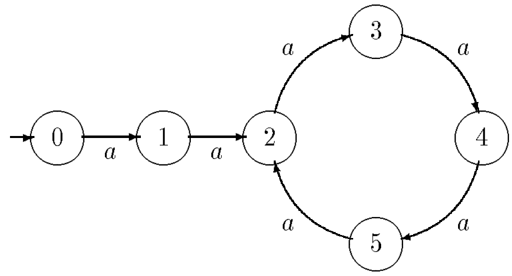
$$n \in (S \cap T) \iff (n \in S \wedge n \in T) \iff (n + pq \in S \wedge n + pq \in T) \iff n + pq \in (S \cap T)$$

Ceci montre que $S \cap T$ est une p.u.p. de $\mathbb{N}^*$. Une récurrence immédiate étend ce résultat à l'intersection d'une famille finie de p.u.p.

• Comme $\mathbb{N}^* \setminus (S \cup T) = (\mathbb{N}^* \setminus S) \cap (\mathbb{N}^* \setminus T)$, les résultats précédents nous montrent que l'intersection de deux p.u.p. de $\mathbb{N}^*$ est elle-même une p.u.p. Par récurrence, le résultat s'étend à une famille finie.

Question 8 Si L est fini, c'est clair : notant $m = \max\limits_{u \in L} |u|$, il suffit de prendre $n_0 = m + 1$ et $p = 1$. Alors $a^{n+p} \in L \iff a^n \in L$ pour tout $n \geqslant n_0$, puisque ces deux assertions sont simultanément fausses !

• Supposons maintenant L infini et considérons un automate fini déterministe complet $\mathcal{A} = (Q, \delta, i, F)$ reconnaissant L. La suite de terme général $\delta^*(i, a^n)$ prend ses valeurs dans l'ensemble fini Q ; il existe donc des exposants n et m distincts tels que $\delta^*(i, a^n) = \delta^*(i, a^m)$. Notons n_0 le plus petit naturel pour lequel il existe $m > n_0$ tel que $\delta^*(i, a_0^n) = \delta^*(i, a^m)$, puis $p = m - n_0$. Le graphe de $\mathcal{A}$ aura l'allure ci-dessous («poêle à frire»). Le dessin correspond au cas $n_0 = 2$, $p = 4$; on n'a pas spécifié quels états étaient finals.



Il est clair que, si $n \geqslant n_0$, alors $\delta^*(i, a^{n+p}) = \delta^*(i, a^n)$, donc $a^{n+p} \in L \iff a^n \in L$, soit encore $n + p \in ||L|| \iff n \in ||L||$. Ceci prouve que $||L||$ est ultimement périodique.

Question 9 • Soit $\mathcal{A} = (Q, \delta, i, F)$ un automate fini déterministe reconnaissant L. Nous allons construire un automate $\mathcal{A}' = (Q', \delta', i, F)$ qui reconnaît $\varphi(L)$. Pour construire δ', nous distinguons pour chaque lettre $x \in X$ trois cas de figure selon la longueur de $\varphi(x)$:

- si $|\varphi(x)| = 1$, on remplace chaque transition de la forme (q, x, q') qui apparaissait dans δ par une transition $\big(q, \varphi(x), q'\big)$;
- si $|\varphi(x)| = 0$, on remplace chaque transition de la forme (q, x, q') qui apparaissait dans δ par une transition instantanée (q, ε, q') ;
- si $|\varphi(x)| > 1$, alors, notant $n = |\varphi(x)|$ et $\varphi(x) = y_1 y_2 \ldots y_n$, on va, pour chaque transition (q, x, q') qui apparaissait dans δ introduire de nouveaux états $q_1, q_2, \ldots, q_{n-1}$ ainsi que les transitions (q, u_1, q_1), (q_1, u_2, q_2) et ainsi de suite jusqu'à (q_{n-1}, u_n, q').

Q' est la réunion de Q et de l'ensemble des nouveaux états ainsi introduits. On vérifie sans peine que $\mathcal{A}'$ reconnaît $\psi(L)$.

• On peut aussi donner une preuve en travaillant sur les expressions rationnelles. Définissons par induction structurelle une application $\widehat{\psi}$ de l'ensemble des expressions rationnelles sur X, sur lui-même, au moyen des règles suivantes :

- $\widehat{\psi}(\emptyset) = \emptyset$
- $\widehat{\psi}(\varepsilon) = \varepsilon$
- $\widehat{\psi}(x) = \psi(x)$ pour tout $x \in X$
- $\widehat{\psi}(e + e') = \widehat{\psi}(e) + \widehat{\psi}(e')$
- $\widehat{\psi}(e \cdot e') = \widehat{\psi}(e) \cdot \widehat{\psi}(e')$
- $\widehat{\psi}(e^*) = \big(\widehat{\psi}(e)\big)^*$

On vérifie aisément que, à une expression rationnelle e décrivant un langage L, $\widehat{\psi}$ associe une expression rationnelle $\widehat{\psi}(e)$ décrivant le langage $\psi(L)$.

Question 10 • Soit ψ le morphisme défini par $\psi(x) = a$ pour toute lettre $x \in \Sigma$. L est rationnel ; donc $\psi(L)$ l'est aussi d'après le résultat de la question 9 ; donc $||\psi(L)||$ est ultimement périodique d'après le résultat de la question 8. Or, comme ψ conserve la longueur, $||L|| = ||\psi(L)||$.

Question 11 • Le langage $L = \{u \in \Sigma^* \mid |u|_a = |u|_b\}$ n'est pas rationnel. $||L||$ est l'ensemble des naturels non nuls pairs, qui est clairement une p.u.p. de $\mathbb{N}^*$.

Question 12 • Soit $\mathcal{A} = (Q, \delta, i, F)$ un automate fini reconnaissant M. L'automate (Q, δ', i, F), où δ' est l'ensemble des triplets (q, x, q') tels que $q \xrightarrow{\psi(x)} q'$ soit un calcul de $\mathcal{A}$, reconnaît $\psi^{-1}(M)$. Remarque : $\mathcal{A}'$ peut contenir des transitions instantanées.

Extractions

Question 13 • Soit $\mathcal{A} = (Q, \delta, i, F)$ un automate fini reconnaissant L. Construisons un automate fini $\mathcal{A}' = (Q', \delta', i', F')$ non déterministe reconnaissant $\varphi_S(L)$. On prendra $Q' = Q \times [\![1, n_0 + 1]\!]$, $i' = (i, 1)$, $F' = F \times [\![1, n_0 + 1]\!]$ et δ' définie comme suit :

- si $1 \leqslant k < n_0$, alors $\delta'\big((q, k), x\big)$ se réduit à $\{\big(\delta(q, x), k + 1\big)\}$;
- si $k = n_0$, alors $q'' \in \delta'\big((q, k), n_0 + 1\big)$ ssi, notant $q' = \delta(q, x)$, il existe une lettre $y \in \Sigma$ telle que $\delta(q', y) = q''$;
- $\delta'\big((q, k), n_0 + 1\big) = \{\big(\delta(q, x), n_0 + 1\big)\}$.

On vérifie que u est reconnu par $\mathcal{A}$ ssi $\varphi_S(u)$ est reconnu par $\mathcal{A}'$: c'est banal si $|u| < n_0$; sinon, les transitions de $\mathcal{A}'$ du deuxième type simulent l'effacement de la lettre en position n_0 dans u.

Question 14 • La question ne se pose que si F n'est pas vide. Notons $n_1, \ldots, n_k$ les éléments de F, rangés par ordre croissant. Notons $S_i = \mathbb{N}^* \setminus \{n_i\}$. Alors φ_S est la composée des k extractions $\varphi_{S_1} \circ \cdots \circ \varphi_{S_k}$, lesquelles conservent la rationalité d'après le résultat de la question précédente. Bien noter l'ordre dans lequel s'effectuent les effacements.

Question 15 • Soit $\mathcal{A} = (Q, \delta, i, F)$ un automate fini reconnaissant L. Soit δ' l'ensemble des triplets (q, x, q') tels qu'il existe un mot u de longueur $p - 1$ tel que $q \xrightarrow{ux}$ soit un calcul de $\mathcal{A}$. L'automate fini (non nécessairement déterministe) $\mathcal{A}' = (Q, \delta', i, F)$ reconnaît $\varphi_S(L)$ (preuve non détaillée ici).

Question 16 • Notons $S = s(\mathbb{N}^*)$. Soient $n_0 \geqslant 1$ et $\pi \geqslant 1$ tels que $n + \pi \in S \iff n \in S$. Notons $F = [\![1, n_0 - 1]\!] \setminus S$, $T = \mathbb{N}^* \setminus F$, et $Q = [\![n_0, n_0 + \pi - 1]\!] \setminus S$. Si Q est vide, on est ramené au cas étudié à la question 14. Sinon, notons $q_1, \ldots, q_k$ les éléments de Q, rangés par ordre croissant. φ_S est la composée $\varphi_T \circ \varphi_1 \cdots \circ \varphi_i$, où φ_k est l'extraction définie comme à la question 15 par $n = q_i$ et $p = \pi - k + i$.

Question 17 • La réponse est négative. Soient $n_0 \geqslant 1$, $p \geqslant 1$ et q un nombre premier au moins égal à n_0. Le nombre $q' = (p + 1)q$ n'est pas premier, et il majore certainement n_0 ; or $q' - q$ est multiple de p. Donc P n'est pas ultimement périodique.

Question 18 • Aucun multiple de 3 supérieur à 3 n'est premier. Donc les a sont tous effacés, à part le premier ; ceci prouve que $\varphi_P(L)$ est contenu dans $M = \{bab^n \mid n \in \mathbb{N}\}$. Montrons l'inclusion inverse. On a $ba = \varphi(bba)$. Soit $n \geqslant 1$; notons p le $(n + 2)$-ième nombre premier, et $q = \lceil p/3 \rceil$. On a $3q - 2 \leqslant p \leqslant 3q$; mais p est premier et $p > 3$, donc $p \neq 3q$. Ainsi, p est égal à $3q - 1$ ou à $3q - 2$, si bien qu'un et seul de ces deux nombres est premier. Ainsi, dans l'intervalle discret $[\![4, 3q]\!]$, il existe exactement n nombres premiers. On en déduit que $\varphi\big((bba)^{q+1}\big) = bab^n$, donc $u \in \varphi(L)$. On conclut en remarquant que M est clairement un langage rationnel.

Question 19 • «Si une partie S de $\mathbb{N}^*$ est telle que $\varphi_S(L)$ soit rationnel quel que soit le langage rationnel L, alors S est ultimement périodique».

Programmation en Caml

Question 20 • Il est clair que si $n_0 = 1$, alors n_0 est minimal. De même, si $n_0 > 1$, et un et un seul des deux naturels $n_0 - 1$ et $n_0 + p - 1$ appartient à G, il n'est pas possible de diminuer n_0.

• Réciproquement, si n_0 est minimal, alors ou bien $n_0 = 1$; ou bien $n_0 > 1$, et alors un et un seul des deux naturels $n_0 - 1$ et $n_0 + p - 1$ appartient à G, sinon on pourrait remplacer n_0 par $n_0 - 1$.

Question 21 • Le plus petit naturel n vérifiant $n > 25 \wedge n \equiv 5\ [8]$ est 29. Donc $G = \{3, 10, 17, 24, 31, 38, 45\} \cup \{29 + 8k \mid k \in \mathbb{N}\}$. On remarque que $29 + 8 \cdot 2 = 45$. Donc $G = \{3, 10, 17, 24, 31, 38\} \cup \{29 + 8k \mid k \in \mathbb{N}\}$. Cette fois, aucun élément n'est compté deux fois ; pour metre en évidence la prépériode, on extrait de la partie périodique les éléments inférieurs à 38, donc 29 et 37 : $G = \{3, 10, 17, 24, 29, 31, 37, 38\} \cup \{45 + 8k \mid k \in \mathbb{N}\}$. La représentation canonique de G est donc $n_0 = 39$, $p = 8$, la prépériode est $\{3, 10, 17, 24, 29, 31, 37, 38\}$ et le motif est $\{45\}$.

Question 22 • Il suffit de transcrire en Caml la définition. Nous aurons besoin d'une fonction qui vérifie qu'une liste ne contient pas de doublon ; nous donnons deux rédactions, l'une très simple s'appuyant sur la fonction `uniq` fournie par l'éoncé, l'autre construite «à la main».

```
let sans_doublon l = (list_length l) = (list_length (uniq l)) ;;

let rec sans_doublon = function
```

```
  | [] -> true
  | t::q -> not (mem t q) & sans_doublon q ;;

let valide g =
 g.n_0 > 0 & sans_doublon g.pre & (forall (fun x -> x < g.n_0) g.pre)
 & g.p > 0 & sans_doublon g.motif
 & (forall (fun x -> x >= g.n_0 & x < g.n_0 + g.p) g.motif) ;;
```

Question 23 • Une p.u.p. est vide ssi sa prépériode et son motif le sont :

```
let est_vide g = g.pre = [] & g.motif = [] ;;
```

Question 24 • Une p.u.p. est finie ssi son motif est vide :

```
let est_finie g = g.motif = [] ;;
```

Question 25 • Les éléments de G sont : les éléments de la prépériode ; et les naturels au moins égaux à n_0, congrus modulo p à un élément y du motif. Soit $n \geqslant n_0$; la condition s'écrit $n = y + qp$ avec $n_0 \leqslant y < n_0 + p$, soit $n_0 \leqslant n - qp < n_0 + p$ ou encore $qp \leqslant n - n_0 < (q+1)p$: donc $q = \lfloor (n - n_0)/p \rfloor$. Alors $y = n - qp$. Traduction immédiate :

```
let appartient n g = if n < g.n_0 then mem n g.pre else
 let q = (n - g.n_0) / g.p in mem (n - q * g.p) g.motif ;;
```

Question 26 • Il s'agit de remplacer la prépériode par son complémentaire dans l'intervalle discret $[\![1, n_0 - 1]\!]$, et le motif par son complémentaire dans l'intervalle discret $[\![n_0, n_0 + p - 1]\!]$. Il est commode d'écrire une fonction auxiliaire `ote_fini_de_discint` pour effectuer ce travail de prise de complémentaire.

```
let ote_fini_de_discint f i j =
 let id = intervalle i j in filtre (fun x -> not (mem x f)) id ;;

let complementaire g =
 {n_0 = g.n_0 ; p = g.p ; pre = ote_fini_de_discint g.pre 1 (g.n_0 - 1);
 motif = ote_fini_de_discint g.motif g.n_0 (g.n_0 + g.p - 1)} ;;
```

Question 27 • Remarquons que la réunion de G et G' est périodique à partir du rang $m = \max(n_0, n'_0)$, une période étant $\pi = \mathrm{ppcm}(p, p')$. La prépériode sera obtenu en filtrant l'intervalle discret $[\![1, m-1]\!]$ par l'appartenance à G ou à G' ; le motif sera obtenu en filtrant de la même manière $[\![m, m + \pi - 1]\!]$.

```
let réunion g g' =
 let bon x = appartient x g or appartient x g'
 let m = max (g.n_0 + g.p) (g'.n_0 + g'.p) and per = ppcm g.p g'.p in
 {n_0 = m ; p = per ;
  pre = filtre bon (intervalle 1 (m - 1)) ;
  motif = filtre bon (intervalle m (m+per-1))
 } ;;
```

Question 28 • Pour l'intersection, nous utilisons la formule $G \cap G' = \overline{\overline{G} \cup \overline{G'}}$.

```
let intersection g g' =
 complementaire ( réunion ( complementaire g) (complementaire g') ) ;;
```

Question 29 Notons G et H les p.u.p. décrites respectivement par **g** et **h**. Il s'agit de vérifier que $n \in H \Rightarrow n \in G$, pour tout $n \geqslant 1$. Notons p^g la période de G, et n_0^g le rang d'entrée dans celle-ci ; définissons de même n_0^h et p^h. Il est clair que la fonction

$$\tau : n \geqslant 1 \mapsto (n \in H) \vee (n \notin G)$$

est ultimement périodique ; sa période est un diviseur de $b = p^g \cdot p^h$, et le rang d'entrée dans sa période est au plus $a = \max(n_0^g, n_0^h)$. Il suffit donc de vérifier que $\tau(n) = \mathrm{VRAI}$ pour $1 \leqslant n < a + b$. On peut diminuer un peu la borne supérieure de l'intervalle à explorer, en notant que la période de τ est au plus égale à $\dfrac{p^g \cdot p^h}{p^g \wedge p^h}$.

```
let contient g h =
 let a = max g.n_0 h.n_0 and b = g.p * h.p / (pgcd g.p h.p) in
 let iv = (intervalle 1 (a+b-1)) in
 forall (fun i -> appartient i g or not(appartient i h)) iv ;;
```

spécifiée comme suit : `contient g h` indique si la p.u.p. décrite par `g` contient la p.u.p. décrite par `h`.

Question 30 • On utilise la caractérisation d'une représentation réduite donnée à la question 20. Si $n_0 = 1$, ou si le motif est vide, la représentation est réduite. Sinon, si $n_0 - 1$ et $n_0 + p - 1$ sont simultanément dans G, ou simultanément hors de G, on peut remplacer n_0 par $n_0 - 1$, mettre à jour la prépériode et le motif en conséquence, et poursuivre la réduction. Le processus se termine en un temps fini, car les valeurs de n_0 forment une suite strictement décroissante de naturels non nuls.

```
let rec ote x = filtre (fun y -> y <> x) ;;

let rec reduction g = match () with
 | () when g.n_0 = 1 or g.motif = [] -> g
 | () when mem (g.n_0 + g.p - 1) g.motif & mem (g.n_0 - 1) g.pre ->
   réduit { n_0 = g.n_0 - 1 ; p = g.p ; pre = ote (g.n_0 - 1) g.pre ;
   motif = (g.n_0 - 1)::ote (g.n_0 + g.p - 1) g.motif }
 | () when not (mem (g.n_0 + g.p - 1) g.motif) & not(mem (g.n_0 - 1) g.pre) ->
   réduit { n_0 = g.n_0 - 1 ; p = g.p ; pre = g.pre ; motif = g.motif }
 | () -> g ;;
```

Question 31 • Une fois la réduction effectuée, il reste à déterminer la plus petite période. On peut bien entendu tester chaque diviseur de p : le coût de cette méthode est un $\mathcal{O}(p\sqrt{p})$. Mais il existe une méthode de coût linéaire, consistant à considérer le motif comme un mot, dont on va déterminer le bord maximal : c'est l'idée qui sous-tend l'algorithme de KNUTH, MORRIS et PRATT. Un mot u est périodique ssi son bord maximal v est tel que $|u| - |v|$ divise $|u|$.

Références bibliographiques

▶ L'idée de base de ce texte (extraction par effacement des lettres dont l'indice n'appartient pas à une partie désignée de $\mathbb{N}^*$) a été prise dans un article (en cours de réaction) d'Armando B. MATOS, intitulé *Regularity preserving letter selections*, trouvé sur le Web à l'url

`http://www.ncc.up.pt/~acm/regn.ps`

▶ L'auteur présente comme question ouverte la réciproque proposée à la question 19.

FIN