Option Informatique en Spé MP et MP*
Devoir surveillé du samedi 9 février 2008 : le corrigé

Problème 1

Question 1 La classe du mot u est un langage fini (elle contient au plus $2^{|u|}$ mots) ; c'est donc un langage rationnel.

Question 2 Il existe 21 classes de mots de longueur 3 ; les voici énumérées par ordre du croissant du représentant : $\{aaa\}$, $\{aab, aba, baa\}$, $\{aac\}$, $\{abb, bab, bba\}$, $\{abc, bac\}$, $\{aca\}$, $\{acb\}$, $\{acc\}$, $\{bbb\}$, $\{bbc\}$, $\{bca\}$, $\{bcb\}$, $\{bcc\}$, $\{caa\}$, $\{cab, cba\}$, $\{cac\}$, $\{cbb\}$, $\{cbc\}$, $\{cca\}$, $\{ccb\}$, $\{ccc\}$.

Question 3 Chaque classe contient au moins un mot ; la classe de a^n contient un seul mot, donc le minimum est égal à 1. Donnons un majorant : il est clair, que pour n fixé, la classe sera d'autant plus grande qu'elle contiendra moins de c ; s'il n'y a aucun c, la classe compte $\binom{n}{|u|_a}$ membres, ce qui est majoré par 2^n.

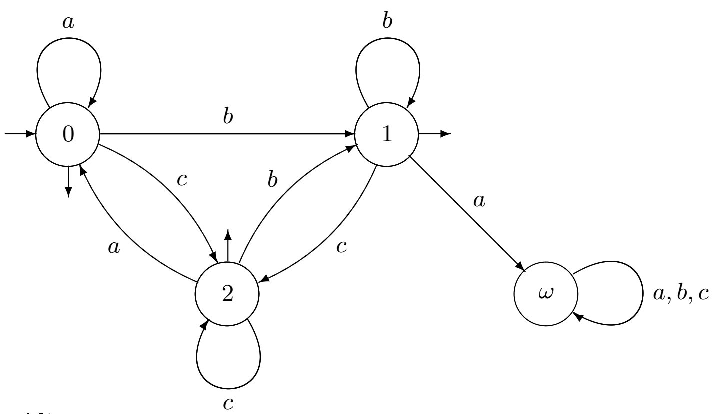
Question 4 Σ^* est infini, mais chaque classe est finie. Donc le nombre de classes est infini.

Question 5 La distance maximale entre deux mots de longueur n sur l'alphabet $\{a,b\}$ est réalisée par les mots $u = a^p b^q$ et $v = b^q a^p$, où $p = \lfloor n/2 \rfloor$ et $q = \lceil n/2 \rceil$. En effet, il s'agit de faire passer chaque b situé à droite du bloc de a, devant ce bloc, et la seule action possible est la permutation de deux lettres consécutives.

Question 6 Ajoutons à l'alphabet Σ une lettre ω, dont nous décidons qu'elle est strictement inférieure à toute les autres. Soient u et v deux mots : quitte à compléter le plus court (disons u) par $k = |v| - |u|$ lettres ω, on se ramène à comparer deux mots $u' = u\omega^k$ et v de même longueur ; il suffit ensuite de noter que $u \preceq v$ ssi $u' \preceq v$. Cette même astuce permet de donner des preuves concises de la réflexivité et de la transitivité de $\preceq$.

Question 7 Chaque représentant est constitué d'un bloc initial de la forme $a^{p_1}b^{q_1}$ (éventuellement vide), suivi de 0, 1 ou plusieurs blocs de la forme $ca^{p_k}b^{q_k}$ (le bloc $a^{p_k}b^{q_k}$ pouvant être vide). L'expression rationnelle $a^*b^*(ca^*b^*)^*$ décrit donc tous les éléments de M ; et il est clair qu'elle ne décrit aucun autre mot.

Question 8 Déduire l'automate demandé de l'expression rationnelle précédente est immédiat.



Question 9 Écriture immédiate :

```
let rec compare_lex = function
 | ([],[]) -> 0
 | ([],_) -> 1
 | (_,[]) -> 2
 | (x::_,y::_) when x<y -> 1
 | (x::_,y::_) when x>y -> 2
 | (_::q,_::r) -> compare_lex(q,r) ;;
```

Question 10 Découpons le mot u en blocs de longueur maximale, ne contenant que des a et/ou des b ; certains de ces blocs peuvent être vides. Deux blocs consécutifs sont séparés par un c. Remplaçons ensuite chaque bloc v par le bloc $a^\alpha b^\beta$, où $\alpha = |v|_a$ et $\beta = |v|_b$. Le mot ainsi obtenu est équivalent à u, et est lexicographiquement minimal dans sa classe.

Question 11 Nous utilisons deux fonctions auxiliaires :

```
let rec f = function
 | (0,0,r) -> r
 | (a,0,r) -> f(a-1,0,A::r)
 | (a,b,r) -> f(a,b-1,B::r) ;;

let rec aux = function
 | (a,b,r,[]) -> f(a,b,r)
 | (a,b,r,A::q) -> aux(a+1,b,r,q)
 | (a,b,r,B::q) -> aux(a,b+1,r,q)
 | (a,b,r,C::q) -> aux(0,0,C::f(a,b,r),q) ;;

let representant m = aux(0,0,[],rev m) ;;
```

Question 12 $\Phi(L) = a^*b$.

Question 13 $\Phi(L) = a^*bab$.

Question 14 Notons que la suite des S_j est croissante pour l'inclusion, et que $\mathbf{D}^\ell(S_j) = S_{j+\ell}$. Considérons le plus petit indice j tel que $q \in S_j$; comme $S_0 = \emptyset$, nous avons certainement $j \geqslant 1$. Notons x la dernière lettre du préfixe $u[1..j]$ de u : il existe une lettre $y < x$ telle que $\delta^*(i, u_1 \dots u_{j-1}y) = q$. Comme $\mathcal{B}$ est déterministe et complet, il peut lire le mot $u_1 \dots u_{j-1}yu_{j+1} \dots u_n$, déduit de u par substitution de la lettre y à la lettre $u_j = x$. Réciproquement, soit v de longueur k, strictement inférieur à $u[1..k]$ pour l'ordre lexicographique. Notons j le plus petit élément de $\in [\![1, k]\!]$ tel que $v_j < u_j$. Notons également $s = \delta^*(i, v[1..j])$; nous avons $u[1..j-1] = v[1..j-1]$, donc $s \in S_j$ mais $s \notin S_{j-1}$. Alors $\mathbf{D}^{k-j}(s) = q$ appartient à S_k.

Question 15 Si $\mathcal{B}$ reconnaît u, alors u est reconnu par $\mathcal{A}$ (et donc appartient à L) : en effet, la projection du calcul de $\mathcal{B}$ sur la première coordonnée nous donne la suite des états par lesquels passe $\mathcal{A}$ lorsqu'il lit u. Enfin, $S_n \cap F = \emptyset$: donc aucun mot de longueur n ne peut être à la fois dans L et être lexicographiquement inférieur à un mot de L. Conclusion : le langage reconnu par $\mathcal{B}$ est l'ensemble des mots de L lexicographiquement minimaux dans leur classe modulo la relation «avoir même longueur que».

Question 16 Soit $\mathcal{A} = (Q, \delta, i, F)$ un automate fini déterministe complet reconnaissant L. L'automate $\mathcal{M}$ suivant reconnaît $\{w\}$:

- ses états sont les triplets (q, γ, S) où $q \in Q$, $\gamma \in \{0,1\}$ et $S \subset Q$;
- son état initial est le triplet $(i, 0, \emptyset)$;
- un état (q, γ, S) de $\mathcal{B}$ est final ssi $q \in F$, $\gamma = 0$ et $S \cap F = \emptyset$;
- sa fonction de transition Δ est définie comme suit : $\Delta\big((q, \gamma, S), x\big) = \big(\delta(q, x), \gamma', S'\big)$ où S' est la réunion de $\mathbf{D}(S)$ et de $\mathbf{T}(S, x)$, et $\gamma' = 1$ si $\gamma = 1$ ou $q \in F$.

L'indicateur γ permet de savoir si l'un des préfixes propres du mot lu appartient à L.

Problème 2

Question 1 • Si $N_i = 0$, alors $N_{i+1} = 0$. Sinon, nous avons $N_{i+1} = \left\lfloor \dfrac{qN_i}{p} \right\rfloor$; mais $q < p$, donc $\dfrac{qN_i}{p} < N_i$ et à plus forte raison $N_{i+1} < N_i$. Dans les deux cas, $N_{i+1} \leqslant N_i$.

Question 2 • Il ne peut exister qu'un nombre fini d'indices i tels que $N_{i+1} < N_i$. Notons k le plus grand d'entre eux. Alors $N_k > 0$ et $N_{k+1} = N_{k+2}$; le résultat de la question précédente montre que ces deux derniers nombres sont égaux à 0, puis que $N_i = 0$ pour tout $i > k$.

Question 3 • Il suffit de remarquer que $a_k = qN_k$, puisque $N_{k+1} = 0$.

Question 4 • $N_0 = 5$, $p = 3$ et $q = 2$. Alors $qN_0 = 10 = 3\times3+1$, donc $a_0 = 1$ et $N_1 = 3$; puis $qN_1 = 6 = 2\times3$, donc $a_1 = 0$ et $N_1 = 2$; puis $qN_2 = 4 = 1 \times 3 + 1$, donc $a_2 = 1$ et $N_2 = 1$; enfin, $qN_3 = 2 = 0 \times 3 + 2$, donc $a_3 = 2$ et $N_4 = 0$. L'algorithme se termine, et la représentation de 5 en base 3/2 est 2101.

Question 5 • Dans le cas particulier $q = 1$, nous retrouvons la numération habituelle en base $p \geqslant 2$.

Question 6 • $N_0 = 17$, $p = 5$ et $q = 3$. Alors $qN_0 = 51 = 10 \times 5 + 1$, donc $a_0 = 1$ et $N_1 = 10$; puis $qN_1 = 30 = 6 \times 5$, donc $a_1 = 0$ et $N_2 = 6$; puis $qN_2 = 18 = 3 \times 5 + 3$, donc $a_2 = 3$ et $N_3 = 3$; puis $qN_3 = 9 = 1 \times 5 + 4$, donc $a_4 = 4$ et $N_4 = 1$; enfin, $qN_4 = 3 = 0 \times 5 + 3$, donc $a_5 = 3$ et $N_5 = 0$. L'algorithme se termine, et la représentation de 17 en base 5/3 est 34301.

Question 7 • L'utilisation du nom `my_n` n'est pas indispensable. La récursivité terminale employée ici a deux avantages : meilleure performance ; et la liste est rendue dans l'ordre attendu, ce qui économise un appel à `rev`.

```
let representation p q my_n =
 let rec aux accu = function
  | 0 -> accu
  | n -> let d = q*n in let (n',a) = (d/p,d mod p) in aux (a::accu) n'
 in aux [] my_n ;;
```

Question 8 • Pour $i \in [\![0, k]\!]$, nous avons $qN_i = pN_{i+1} + a_i$. En multipliant les deux membres par p^i/q^{i+1}, nous obtenons la relation $\dfrac{p^i N_i}{q^i} = \dfrac{p^{i+1} N_{i+1}}{q^{i+1}} + \dfrac{p^i a_i}{q^{i+1}}$, laquelle se prête au télescopage ; comme $N_0 = n$ et $N_{k+1} = 0$, il vient $n = \sum\limits_{0 \leqslant i \leqslant k} \dfrac{p^i a_i}{q^{i+1}}$. Donc la suite de terme général $u_i = p^i/q^{i+1}$ répond à la question.

Question 9 • Notez le `rev` rendu nécessaire par l'ordre de lecture de la liste.

```
let conversion p q l =
 let rec aux = function
  | [] -> 0
  | t::r -> (p*(conversion p q r)+t)/q
 in aux (rev l) ;;
```

Question 10 • Soit w appartenant à $L_{p/q}$, de longueur $k \geqslant 2$. Il suffit de montrer que le préfixe w' de w, de longueur $k-1$, est encore dans $L_{p/q}$. Notons n l'entier représenté par w en base p/q, et $w = w_1 \dots w_k$; d'après la première étape de l'algorithme, nous avons $N_0 = n$ et $qN_0 = pN_1 + w_k$, donc $w_1 \dots w_{k-1}$ appartient à $L_{p/q}$, car il représente $\dfrac{qN_0 - w_k}{p}$ en base p/q.

Question 11 • Soit w appartenant à $L_{p/q}$, de longueur $k \geqslant 1$. w représente un naturel non nul n. Notons $n' = \lceil pn/q \rceil$; nous aurons $pn/q \leqslant n' < pn/q + 1$, soit $0 \leqslant n'q - pn < q$. Notons $w_{k+1} = n'q - pn$: alors $w_{k+1} \in [\![0, q-1]\!]$, donc $w_{k+1} \in \Sigma$; w est préfixe du mot $w_1 \dots w_k w_{k+1}$, lequel appartient à $L_{p/q}$, car il représente le naturel n' en base p/q. Notons qu'il peut exister plusieurs solutions : par exemple, en base 3/2, le mot 21011 représente 8 ; il est préfixe de 210110, qui représente 12, et de 210112, qui représente 13.

Question 12 • Solution de facilité : $a_k \neq 0$, donc aucun mot de $L_{3/2}$ ne commence par un 0 ; ainsi, le langage rationnel infini `0(0+1+2)*` est disjoint de $L_{3/2}$. Solution plus intéressante : il n'existe aucun mot de $L_{3/2}$ commençant par 22 : en effet, si $a_k = a_{k-1} = 2$, alors $2N_k = a_k$ implique $N_k = 1$; puis $2N_{k-1} = 3N_k + a_{k-1}$ implique $2N_{k-1} = 5$, ce qui est contradictoire. Donc le langage rationnel infini `22(0+1+2)*` est disjoint de $L_{3/2}$.

FIN