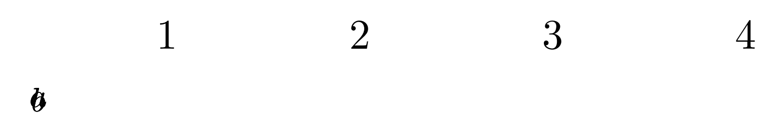
Ilog3 (Théorie des Langages): Devoir de Décembre 2008.

Il est recommandé de rédiger lisiblement et avec soin sa copie. En particulier, les réponses doivent être justifiées, il est conseillé de traiter en priorité les questions que l'on sait faire en en indiquant clairement la référence.

I) Question de Cours & Exercices:

1) a) Qu'est-ce qu'un langage? Donnez des exemples significatifs.
Dans cette question de cours, à partir de maintenant, $A = \{a,b\}$.
b) Comment est défini le langage des mots sans facteur b^3 (on notera ce dernier L_{b^3}).
c) Montrer, en analysant les chemins que l'automate suivant reconnait L_{b^3}.

2) a) Écrire les éléments matriciels de l'automate ci-dessous $(I,M(a),M(b),F)$).



b) Quel est le langage reconnu par cet automate?
c) Peut-on réduire son nombre d'états de cet automate?
3) (Généralisation) L'automate du (2) est le 4ème d'une suite d'automates $\mathcal{A}_n$ définis comme suit $Q = [1..n]$ (entiers compris entre 1 et n); états initial et final 1; transitions $j.a = j+1$ si $j < n$ et $j.b = j-1$ si $j > 1$. Quel est le langage L_n reconnu par $\mathcal{A}_n$?

II) Exercices

A) **Sous-mots et mélanges (shuffle)**

1) Soit $w = a_1 a_2 \cdots a_n$ et $\{i_1,i_2,\cdots i_k\} = I \subset [1..n]$ avec $i_1 < i_2 < \cdots < i_k$. On note $w[I]$ le mot $a_{i_1} a_{i_2} \cdots a_{i_k}$. On l'appelle un sous-mot de longueur k.
a) Pour chacun des mots suivants donner la liste des sous-mots de longueur k. Pour vérification, le nombre de sous-mots distincts de longueur $1,2,\cdots |w|$ est donné entre parenthèses.

i) $w = aba$ (2,3,1) ii) abc (3,3,1) iii) $abab$ (2,4,4,1).

2) Soient u,v deux mots, on dit que w de longueur n est un mélamge (shuffle) de u (de longueur p) et v (de longueur q) si il existe une partition $I + J = [1..n]$ (c'est à dire $I \cup J = [1..n]$ et $I \cap J = \emptyset$) telle que $w[I] = u$; $w[J] = v$. On notera $u \sqcup\!\sqcup v$ le langage formé des shuffles de u et v. Par exemple

$$abc \sqcup\!\sqcup xy = abcxy + abxcy + abxyc + axbcy + axbyc + axybc + xabcy + xabyc + xaybc + xyabc$$

a) Montrer que l'opération $\sqcup\!\sqcup$ est entièrement définie par

$$u \sqcup\!\sqcup \epsilon = \epsilon \sqcup\!\sqcup u = u \ ; \ au \sqcup\!\sqcup bv = a(u \sqcup\!\sqcup bv) + b(au \sqcup\!\sqcup v) \tag{1}$$

b) Pour deux langages $L,M \in A^*$, on note $L \sqcup\!\sqcup M$ le langage

$$L \sqcup\!\sqcup M = \sum_{u \in L,\ v \in M} u \sqcup\!\sqcup v \tag{2}$$

montrer que, pour toute lettre $a \in A$

$$a^{-1}(L \sqcup\!\sqcup M) = (a^{-1}L) \sqcup\!\sqcup M + L \sqcup\!\sqcup (a^{-1}M) \tag{3}$$

c) en déduire que, si L,M sont rationnels alors $L \sqcup\!\sqcup M$ aussi.

B) **Infiltration**

1) Soient u,v deux mots, on dit que w de longueur n est une infiltration de u (de longueur p) et v (de longueur q) si il existe un recouvrement I,J de $[1..n]$ (c'est à dire $I \cup J = [1..n]$) tel que $w[I] = u$; $w[J] = v$. On notera $u \uparrow v$ le langage formé des infiltrations de u et v. Par exemple

$$ab \uparrow ac = abac + aabc + aacb + acab + abc + acb$$

a) Montrer que l'opération $\uparrow$ est entièrement définie par

$$u \uparrow \epsilon = \epsilon \uparrow u = u \ ; \ au \uparrow bv = a(u \uparrow bv) + b(au \uparrow v) + \delta_{a,b}(u \uparrow v) \tag{4}$$

où $\delta_{a,b} = 0$ si $a \neq b$ (c'est à dire que le facteur n'y est pas) et 1 si $a = b$ (c'est à dire que le facteur y est).

b) Pour deux langages $L,M \in A^*$, on note $L \uparrow M$ le langage

$$L \uparrow M = \sum_{u \in L,\ v \in M} u \uparrow v \tag{5}$$

montrer que, pour toute lettre $a \in A$

$$a^{-1}(L \uparrow M) = (a^{-1}L) \uparrow M + L \uparrow (a^{-1}M) + (a^{-1}L) \uparrow (a^{-1}M) \tag{6}$$

c) en déduire que, si L,M sont rationnels alors $L \uparrow M$ aussi.

III. (Petit) Problème

Le but du problème ci-dessous est de définir une classe de mots (les grands Dycks) de deux manières (par une condition numérique, par une grammaire) et de donner l'équivalence entre ces deux définitions.

1) Soit $A = \{x,\bar{x}\}$, un alphabet à deux lettres. On définit une fonction de poids $\pi : A^* \mapsto \mathbb{Z}$ par

$$\pi(x) = 1,\ \pi(\bar{x}) = -1,\ \pi(uv) = \pi(u) + \pi(v) \tag{7}$$

a) Montrer, par récurrence sur la longueur du mot w, que $\pi(w) = |w|_x - |w|_{\bar{x}}$. En déduire que si $\pi(w) = 0$ alors $|w|$ est pair.

On définit les langages suivants :

$$GD := \{w \in A^* \big| \pi(w) = 0\} \tag{8}$$
$$GD_+ := \{w \in GD \big| w = uv \Longrightarrow \pi(u) \geq 0\} \tag{9}$$
$$GD_- := \{w \in GD \big| w = uv \Longrightarrow \pi(u) \leq 0\} \tag{10}$$

b) En associant a chaque lettre un "pas" (nord-est pour x et sud-est pour $\bar{x}$), on fait correspondre à chaque mot un chemin. Donner la signification graphique de chacun des langages ci-dessus.

c) Voici, pour vérification, les distributions de longueur de GD,GD_+ (la distribution de GD_- est la même que celle de GD_+).

n	0	2	4	6	8
$\lvert GD \cap A^n \rvert$	1	2	6	20	70
$\lvert GD_+ \cap A^n \rvert$	1	1	2	5	14

donner la liste des mots de GD de longueur 2,4 et de GD_+ de longueur 6,8.

c) Montrer que GD_+ est le langage défini par la grammaire suivante

$$T \to \epsilon;\ T \to xT\bar{x};\ T \to TT \tag{11}$$

2) a) Donner une description équivalente à (11) pour GD_-.

b) En déduire que GD est le langage engendré par la grammaire suivante

$$P \to \epsilon;\ P \to xP\bar{x};\ P \to PP;\ N \to \epsilon;\ N \to \bar{x}Nx;\ N \to NN;$$
$$T \to \epsilon;\ T \to TT;\ T \to xP\bar{x};\ T \to \bar{x}Nx$$