Corrigé TD n°3

Partie II - Automates

II.B - $\mathcal{A}$ reconnaît seulement 4 mots : *aacaa*, *abcab*, *bacba*, *bbcbb*. Donc $\mathrm{L}(\mathcal{A}) = L_2 = \{u.c.u,\ u \in \{a,b\}^* \ /\ |u| = 2\}$.

II.C - $L_n = \{u.c.u,\ u \in \{a,b\}^* \ /\ |u| = n\}$.

<u>Notons P_n l'ensemble des mots sur $\{a,b\}$ de longueur $\leq n$.</u>

Sachant qu'il y a un mot de longueur nulle (ε), 2 mots de longueur 1, ..., 2^k mots de longueur k, on trouve que P_n a $N = \sum_{k=0}^{n} 2^k = 2^{n+1} - 1$ éléments.

Sur l'alphabet $\Sigma = \{a,b,e,f\}$, notons Q l'ensemble des mots de la forme $e.u$ ou $u.f$ avec $u \in P_n$.

Il est clair que $\boxed{|Q| = 2\,N = 2^{n+2} - 2.}$

Considérons l'automate déterministe $\mathcal{A}_n = \langle Q, e, \{f\}, A, \delta\rangle$ où δ est une <u>fonction</u> de $Q \times A$ dans Q définie de la façon suivante :

- ⋆ si s est un état de la forme $e.u$ avec $|u| < n$, alors $\delta(s,a) = s.a$ et $\delta(s,b) = s.b$.
- ⋆ si $s = e.u$ avec $|u| = n$, alors $\delta(s,c) = u.f$.
- ⋆ si $s = u.f$ avec $|u| \neq \varepsilon$, de la forme $s = u_1u_2\dots u_kf$ où les u_i sont des lettres a ou b, alors $\delta(s,u_1) = u_2\dots u_nf$.
- ⋆ δ n'est pas définie sur les couples non évoqués.

Montrons que cet automate reconnaît le langage L_n.

<u>$L_n \subset \mathrm{L}(\mathcal{A}_n)$</u> : en effet, soit $m \in L_n$. Il est de la forme $m = u_1u_2\dots u_ncu_1u_2\dots u_n$ avec les $u_i \in \{a,b\}$ et on vérifie facilement par récurrence que:

- $\forall\, k \in [\![1;n]\!]$, $\delta^*(e, u_1\dots u_k) = eu_1\dots u_k$,
- $\delta^*(e, u_1\dots u_nc) = u_1\dots u_nf$.
- $\forall\, k \in [\![1;n-1]\!]$, $\delta^*(e, u_1\dots u_ncu_1\dots u_k = u_{k+1}\dots u_nf$.
- $\delta^*(e,m) = f$, ce qui assure que $m \in \mathrm{L}(\mathcal{A}_n)$.

<u>$\mathrm{L}(\mathcal{A}_n) \subset L_n$</u> : en effet, soit $m \in \mathrm{L}(\mathcal{A}_n)$.

Ainsi $\delta^*(e,m)$ est <u>défini</u> et est égal à f. En particulier, $\delta^*(e,p)$ est défini pour tout préfixe p de m.

- Puisque $\delta^*(e,m) = f$, nécessairement m a pour longueur $2\,n+1$.
  m peut donc s'écrire $m = u.x.w$ avec $u \in A^*$, $|u| = n$, $x \in A$, $w \in A^*$, $|w| = n$.
- Puisque $\delta^*(e,u)$ est défini, nécessairement $u \in \{a,b\}^*$.
- Puisque $\delta^*(e,u.x)$ est défini, nécessairement $x = c$.
- Puisque $\delta^*(e,u.x.w)$ est défini, nécessairement $u = w$. Donc $m \in L_n$.

II.D - $L'_n = \{v_1.u.v_2.c.w_1.u.w_2,\ u \in \{a,b\},\ |u = n|\}$.

On reprend le même ensemble d'états Q qu'à la question précédente (de cardinal $2^{n+2} - 2$)

Considérons l'automate non déterministe $\mathcal{A}'_n = \langle Q, e, \{f\}, A, \delta\rangle$ où cette fois δ est une <u>application</u> de $E \times A$ dans $\mathcal{P}(E)$ définie de la façon suivante (obtenue en rajoutant des transitions à l'automate $\mathcal{A}_n$:

- ⋆ $\delta(e,a) = \{e, ea\}$, $\delta(e,b) = \{e, eb\}$, $\delta(e,c) = \{e\}$.
- ⋆ si s est un état de la forme $e.u$ avec $|u| < n$, alors $\delta(s,a) = \{s.a\}$, $\delta(s,b) = \{s.b\}$ et $\delta(s,c) = \emptyset$.
- ⋆ si s est un état de la forme $e.u$ avec $|u| = n$, alors $\delta(s,a) = \{s\}$, $\delta(s,b) = \{s\}$, $\delta(s,c) = \{s, u.f\}$.
- ⋆ si $s = f.u$ avec $|u| = n$ et $u = u_1\dots u_n$, alors $\delta(s,u_1) = \{s, u_2\dots u_nf\}$, $\delta(s,c) = \{s\}$.
- ⋆ si $s = u.f$ avec $|u| < n$, de la forme $s = u_1u_2\dots u_kf$, alors $\delta(s,u_1) = \{u_2\dots u_nf\}$ et $\delta(s,c) = \emptyset$.
- ⋆ Pour $x \in A$, $\delta(f,x) = \{f\}$.

On démontre de façon semblable, mais plus compliquée qu'au **II.C** que $\mathcal{A}'_n$ reconnaît L'_n.

II.E - Graphe d'un automate reconnaissant L'_2.

On l'obtient en rajoutant, dans le graphe fourni par l'énoncé, les transitions $s \xrightarrow{a,b,c} s$ partant de s et revenant à s pour tous les états s <u>sauf</u> pour $e_a,\ e_b,\ f_a,\ f_b$ (*figure non reproduite dans ce corrigé*).

II.F - <u>Remarque</u> : les mots de L'_n sont de longueur $\geq 2n+1$ et ceux de longueur $2n+1$ sont ceux de L_n.

Si $m \in L'_n$ et $|m| = 2\,n+1$, alors nécessairement $m \in L_n$ de la forme $u_1 \dots u_n c u_1 \dots u_n$.

Considérons une suite acceptante pour m : $s_0 = e \xrightarrow{u_1} s_1 \xrightarrow{u_2} s_2 \dots s_n \xrightarrow{c} s_{n+1} \xrightarrow{u_1} s_{n+2} \dots \xrightarrow{u_n} s_{2n+1} \in F$.

Raisonnons par l'absurde en supposant qu'il existe deux états égaux s_i et s_j avec $i < j$.

- si $0 \leq i < j \leq n$, alors l'automate $\mathcal{A}$ reconnaît aussi le mot $u_1 \dots u_i u_{j+1} \dots u_n c u_1 \dots u_n$ qui n'est pas dans L'_n car de
longueur $< 2n+1$.
- si $n+1 \leq i < j \leq 2n+1$, alors l'automate $\mathcal{A}$ reconnaît aussi le mot $u_1 \dots u_n c u_1 \dots u_i u_{j+1} \dots u_n$ qui n'est pas dans L'_n.
- si $i \leq n < n+1 \leq j$, alors l'automate $\mathcal{A}$ reconnaît aussi le mot $u_1 \dots u_i u_{j+1} \dots u_n$ qui n'est pas dans L'_n car ne contient pas c par exemple.

Ainsi tous les états s_i sont <u>distincts</u>.

II.G - Notons G_1 l'ensemble des états accessibles par des mots de P_n. Il suffit de montrer que Card $(G_1) \geq 2^{n+1} - 1 =$Card (P_n).

Chaque fois qu'on prend un mot $u \in P_n$, on peut le considérer comme le préfixe d'un mot $m = u.w.c.u.w \in L_n$. Il existe une suite acceptante pour m et donc un état x accessible par u. De même, si $v \in P_n$, il existe un état y accessible par v.

Montrons que si $u \neq v$, alors $x \neq y$. Il en résultera bien que Card $(G_1) \geq$Card (P_n).

Raisonnons par l'absurde en supposant $u \neq v$ et $x = y$.

Il existe une suite acceptante $s_0 \dots s_k s_{k+1} \dots s_{2n+1}$ avec $s_k = x$ et $k = |u|$ pour un mot m de la forme $u.w.c.u.w \in L'_n$.

De même, il existe une suite acceptante $s'_0 \dots s'_l s_{l+1} \dots s'_{2n+1}$ avec $s'_l = y$ et $l = |v|$.

On a par exemple $l \leq k$. On constate qu'alors $s'_0 \dots s'_l s_{k+1} \dots s_{2n+1}$ est une suite acceptante pour un mot $v.w.c.u.w$ de longueur $\leq 2n+1$ qui n'est pas dans L'_n car $u \neq v$, d'où une contradiction.

II.H - Considérons l'ensemble G_2 des états de $\mathcal{A}$ accessibles par des mots de la forme $u.c.u'$ avec $|u| = n$ et u' préfixe de u.

En écrivant $u = u'.u''$, cela revient à considérer les états «co-accessibles» par des mots u'' de P_n.

On montrerait de même que Card $(G_2) \geq 2^{n+1} - 1$.

Vérifions par l'absurde que $G_1 \cap G_2 = \emptyset$.

Soit $s \in G_1 \cap G_2$: s est accessible par un mot $u \in P_n$ et co-accessible par un mot $v \in P_n$.

Il en résulte que le mot uv est reconnu par l'automate considéré. Or il ne contient pas c, donc il n'est pas dans L'_n, d'où une contradiction.

<u>En résumé</u> : le nombre d'états de $\mathcal{A}$ est $\geq$Card $(G_1 \cup G_2) =$ Card $(G_1)+$ Card $(G_2) \geq 2\,(2^{n+1} - 1) = 2^{2n+2} - 2$.

II.I - Une expression rationnelle de L'_1 est : $\boxed{(a+b)^*\big(aa\mathcal{E}aa + ab\mathcal{E}ab + ba\mathcal{E}ba + bb\mathcal{E}bb\big)(a+b)^*}$ avec $\mathcal{E} = (a+b)^* c (a+b)^*$.

II.J - L'_1 est l'ensemble des mots sur $A = \{a, b, c\}$ qui contiennent une et une seule occurrence de la lettre c et dans lesquels on a soit : $\dots a \dots c \dots a \dots$, soit : $\dots b \dots c \dots b \dots$.

Les mots de $A - L'_1$ sont les mots :

- ⋆ qui ne contiennent pas la lettre c : $\mathcal{E}_1 = (a+b)^*$.
- ⋆ ou qui contiennent au moins deux lettres c : $\mathcal{E}_2 = (a+b)^* c (a+b+c)^* c (a+b)^*$.
- ⋆ ou qui contiennent une seule lettre c qui est
  - précédée d'au moins un a, mais pas suivie de a : $\mathcal{E}_3 = b^* a (a+b)^* c b b^*$.
  - précédée d'au moins un b, mais pas suivie de b : $\mathcal{E}_4 = a^* b (a+b)^* c a a^*$.
  - suivie d'au moins un a, mais pas précédée de a : $\mathcal{E}_5 = b^* b c (a+b)^* a b^*$.
  - suivie d'au moins un b, mais pas précédée de b : $\mathcal{E}_6 = a^* a c (a+b)^* b a^*$.

Une expression rationnelle de $A - L'_1$ est : $\underline{\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3 + \mathcal{E}_4 + \mathcal{E}_5 + \mathcal{E}_6}$.