TD 6

Exercice 1. *Petite liste*

Parmi les langages suivants il y a *au moins un* langage algébrique et *au moins un* non-algébrique. Choisissez deux langages tels que un est algébrique et l'autre non et démontrez le (pour montrer qu'un langage est non-algébrique vous utiliserez le lemme de l'étoile pour les langages hors-contexte) :

1. $\{w \in \{a,b\}^*$ tels que $|w|_b = 2|w|_a + 3\}$
2. $\{w\#x \mid w, x \in \{a,b\}^*$ et w est un sous-mot de $x\}$.
3. $\{a^p \mid p$ est premier$\}$.
4. $\{a^{n_0}ba^{n_1}b\cdots a^{n_k}b | k \geq 0$ et $\exists j \geq 0, n_j \neq j\}$

Exercice 2. *Mélange*

Soit Σ un alphabet fini. Soient u et v deux mots sur Σ^*. On appelle mélange des mots u et v, et l'on note $\mathrm{Mel}(u,v)$ l'ensemble des mots de Σ^* défini par :
— si $u = \varepsilon$, $\mathrm{Mel}(u,v) = \{v\}$
— si $v = \varepsilon$, $\mathrm{Mel}(u,v) = \{u\}$
— si $u = xu'$ et $v = yv'$ avec $x, y \in \Sigma$, $\mathrm{Mel}(u,v) = x.\mathrm{Mel}(u',v) \cup y.\mathrm{Mel}(u,v')$.

Si L et L' sont deux langages, on définit $\mathrm{Mel}(L,L') = \bigcup_{u\in L, v\in L'} \mathrm{Mel}(u,v)$.

1. On considère les langages $L = (aa)^*$ et $L' = (bbb)^*$. Montrer que $\mathrm{Mel}(L,L')$ est rationnel.
2. Le mélange de deux langages rationnels est-il toujours rationnel ?
3. On considère $L = \{a^nb^n \mid n \geq 0\}$ et $L' = c^*$. Montrer que $\mathrm{Mel}(L,L')$ est algébrique.
4. Montrer que le mélange d'un langage rationnel et d'un langage algébrique est algébrique.
5. (Bonus) Qu'en est-il du mélange de deux langages algébriques ?

Exercice 3. *Morceaux de grammaires*

Donner des grammaires algébriques engendrant les langages suivants.

1. L'ensemble des palindromes sur $\{a,b\}$ et son complémentaire.
2. L'ensemble des mots sur $\{a,b\}$ de longueur impaire.
3. L'ensemble des mots sur $\{a,b\}$ ayant le même nombre d'occurrences de a que de b.
4. L'ensemble des mots sur $\{a,b\}$ ayant deux fois plus de a que de b.
5. $\{w\#\overline{w}\#,\ w \in (a+b)^*\}$, avec $\overline{w_1w_2\ldots w_n} = w_n\ldots w_2w_1$.
6. $\{w\#w' | w, w' \in (a+b)^*$ et $w \neq w'\}$.
7. L'ensemble des mots de $(a+b)^*$ qui ne sont pas de la forme ww.
   Indication : les mots qui ne sont pas de la forme ww et qui sont de longueur paire sont de la forme xy avec x et y de longueur impaire, et une autre condition sur x et y.

Exercice 4. *puis*

Pour cet exercice, on considère que pour un langage X et $n \in \mathbb{N}, X^n = X \cdot X \cdot \ldots \cdot X = \{w_1 w_2 \ldots w_n \mid \forall i, w_i \in X\}$.

1. Soit X et Y deux langages réguliers. Montrer que $\bigcup_{n\in\mathbb{N}} X^n \cap Y^n$ est algébrique mais pas forcément régulier.
2. Trouver trois langages X, Y et Z tels que $\bigcup_{n\in\mathbb{N}} X^n \cap Y^n \cap Z^n$ n'est pas algébrique.

Exercice 5. *2*

Définition 1. *Soit $G = (V, \Sigma, P, S)$ une grammaire hors-contexte. La grammaire est dite* autoenchâssante[1] *s'il existe $X \in V$ tel que $X \overset{+}{\Rightarrow} \alpha X \beta$ avec $\alpha, \beta \in (V \cup \Sigma)^+$ ($\alpha, \beta \neq \epsilon$).*
Un langage hors-contexte L est dit autoenchâssant *si toute grammaire hors-contexte le générant est autoenchâssante.*

On souhaite montrer le théorème suivant :

Théorème 1. *Un langage hors-contexte infini est régulier si et seulement s'il n'est pas autoenchâssant.*

1. Montrer l'implication directe.

On veut montrer l'implication inverse. Soit donc un langage L hors-contexte non autoenchâssant. Soit $G = (V, \Sigma, P, S)$ une grammaire hors-contexte le générant telle que tout symbole non terminal est *accessible*[2] et sans règle de la forme $X \to Y$, $X, Y \in V$.

2. Supposons que pour tout X, il existe une dérivation $X \overset{*}{\Rightarrow} \gamma S \gamma'$. Montrer qu'alors, G est nécessairement linéaire à gauche ou à droite.
3. On suppose maintenant qu'il existe $X \in V$ tel que $X \overset{*}{\not\Rightarrow} \gamma S \gamma'$. Montrer par induction sur $|V|$ que L est rationnel.

Exercice 6. *One direction*

Soit Σ un alphabet et $\# \notin \Sigma$ un caractère distingué. Posons $\Sigma_\# = \Sigma \cup \{\#\}$. Appelons automate bidirectionel non-déterministe sur Σ un tuple $A = (Q, I, \Delta, F)$ où Q est un ensemble, $I, F \subseteq Q$ et $\Delta \subseteq Q \times \Sigma_\# \times \{Gauche, Droite\} \times Q$. Une configuration de l'automate est un élément $(u, q, v) \in \Sigma_\#^* \times Q \times \Sigma_\#^*$. On définit la relation de transition $\to$ sur les configurations.

$$\begin{array}{lcll} (u,q,av) & \to & (ua,r,v) & \text{si } (q,a,Droite,r) \in \Delta \\ (ua,q,bv) & \to & (u,r,abv) & \text{si } (q,b,Gauche,r) \in \Delta \end{array}$$

Un mot $w \in \Sigma^*$ est reconnu par A si et seulement si il existe $q_0 \in I$, $q_f \in F$ et des mots $u, v \in \Sigma_\#^*$ tels que $(\epsilon, q_0, \#w\#) \to^* (u, q_f, v)$.
Montrez que la classe des langages reconnus par ces automates est exactement la classe des langages rationnels.

1. Ou *self-embedding* en anglais.
2. Pour tout $X \in V$, il existe une dérivation $S \overset{*}{\Rightarrow} \alpha X \beta$.