## Feuille de TD n° 6 : Expressions régulières et langages reconnaissables

**Exercice 1 :**

- Soit $E_1 = (a+b)^*$. Les mots $a^2$, $a^2b$ et $ba^2$ appartiennent-ils au langage décrit par $E_1$ ?
- Soit $E_2 = a^*b^*c^*$. Les mots $a^2$, $ab$, $a^2b$, $cb$ et $aabc$ appartiennent-ils au langage décrit par $E_2$ ?
- Soit $E_3 = (a+b)^*abb$. Les mots $a^2$, $a^2b$, $ab^2ab^2$ et $a^3b^2$ appartiennent-ils au langage décrit par $E_3$ ?

**Exercice 2 :**

Décrire (en français) les langages définis par les expressions régulières ci-dessous, puis donner des automates les reconnaissant.

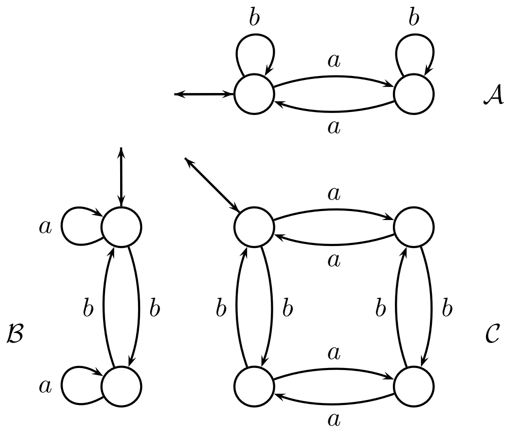
- $E_1 = (aa+b)^*(a+bb)^*$,
- $E_2 = (a+ba+bba)^*(\varepsilon+b+bb)$,
- $E_3 = (aa+bb+(ab+ba)(aa+bb)^*(ab+ba))^*$.

**Exercice 3 :**

Donner des expressions régulières décrivant les langages ci-dessous :

- $\mathcal{L}_1 = \{u \in A^* :$ toute occurrence de $b$ est immédiatement suivie de deux occurrences de $a\}$,
- $\mathcal{L}_2 = \{u \in A^* : u$ ne contient pas deux $a$ successifs$\}$,
- $\mathcal{L}_3 = \{u \in A^* :$ le nombre d'occurrences de $a$ dans $u$ est pair$\}$,
- $\mathcal{L}_4 = \{u \in A^* :$ les blocs de $a$ dans $u$ sont alternativement de longueur paire et impaire$\}$.

**Exercice 4 : intersection de langages et automate produit**




1. Décrire les langages $\mathcal{L}(\mathcal{A})$ et $\mathcal{L}(\mathcal{B})$ reconnus respectivement par les automates $\mathcal{A}$ et $\mathcal{B}$.
2. Après avoir lu un nombre pair de $b$, dans quels états peut se trouver l'automate $\mathcal{C}$ ?
3. Décrire le langage $\mathcal{L}(\mathcal{C})$ reconnu par l'automate $\mathcal{C}$ et conclure.
4. Que faut-il changer à $\mathcal{C}$ pour que $\mathcal{L}(\mathcal{C}) = \mathcal{L}(\mathcal{A}) \cup \mathcal{L}(\mathcal{B})$ ?

**Exercice 5 : applications du lemme de l'étoile**

Soit $A = \{a, b\}$

Les langages suivants sont -ils reconnaissables par un automate ?

- $\mathbf{L}_1 = \{a^nb^n \mid n \geq 0\}$
- $\mathbf{L}_2 = \{u \in A^* \mid |u|_a = |u|_b\}$
- $\mathbf{L}_3 = \{u \in A^* \mid |u|_a < |u|_b\}$

- $\mathbf{L}_4 = \{baba^2ba^3 \dots ba^n \mid n \geq 0\}$
- $\mathbf{L}_5 = \mathbf{L}_2 \setminus A^*\{aa, bb\}A^*$
- $\mathbf{L}_6 = \{u \in A^* \mid \text{u est un palindrome}\}$
- $\mathbf{L}_7 = \{uu \mid u \in A^*\}$
- $\mathbf{L}_8 = FG(\mathbf{L}_1)$
- $\mathbf{L}_9 = FG(\mathbf{L}_2)$
- $\mathbf{L}_{10} = \{a^n \mid n \text{ est un nombre premier}\}$

où $FG(\mathcal{L})$ est l'ensemble des facteurs gauches (ou préfixes) des mots du langage $\mathcal{L}$. On en profitera pour montrer que si $\mathcal{L}$ est reconnaissable alors $FG(\mathcal{L})$ est reconnaissable.

**Exercice 6 : petite variation...**

On considère ici l'alphabet $B = \{a, b, c\}$.

On dit qu'un mot contient un carré si on peut l'écrire $uvvw$ avec $v \neq \varepsilon$. On admettra (c'est loin d'être évident) qu'on peut construire des mots sans carré arbitrairement longs sur l'alphabet $\{a, b, c\}$. Montrer que le langage suivant n'est pas reconnaissable par un automate déterministe *complet* (dans la suite du cours, on verra que cela implique qu'il n'est reconnaissable par aucun automate fini) :

$\mathcal{K} = \{u \in B \mid u \text{ contient un carré}\}$