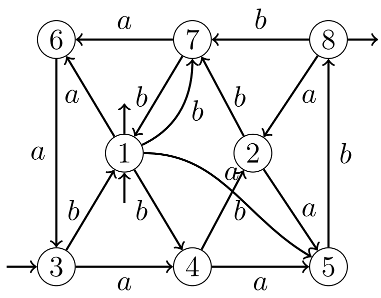
Solution - TD 4 - Clôture des langages reconnaissables, expressions régulières

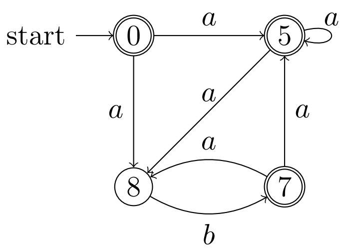
Licence 3 - Université de Bordeaux

Solution de l'exercice 1 :

1. On donne un automate équivalent sans ϵ-transitions à $\mathcal{A}_1$:



On donne un automate équivalent sans ϵ-transitions à $\mathcal{A}_2$:



2. On construit un automate pour $L(\mathcal{A}_1) \cup L(\mathcal{A}_2)$:

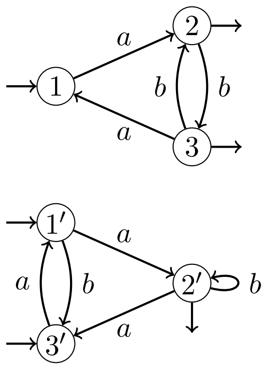


FIGURE 1 – Automate pour $L(\mathcal{A}_1) \cup L(\mathcal{A}_2)$

On calcule l'automate produit pour $L(\mathcal{A}_1) \cap L(\mathcal{A}_2)$:

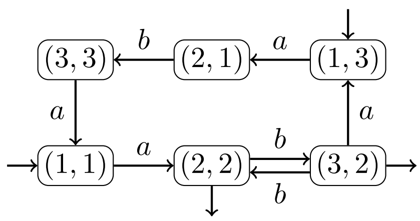


FIGURE 2 – Automate pour $L(\mathcal{A}_1) \cap L(\mathcal{A}_2)$

Pour $L(\mathcal{A}_1) \bullet L(\mathcal{A}_2)$, on commence par faire une version avec ϵ-transitions :

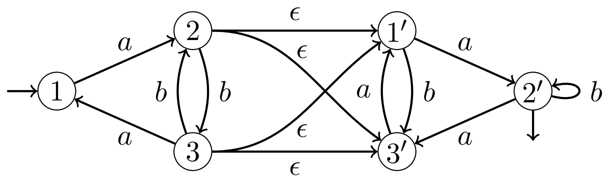


FIGURE 3 – Automate pour $L(\mathcal{A}_1) \bullet L(\mathcal{A}_2)$ (avec ϵ-transitions)

On élimine ensuite les ϵ-transitions :

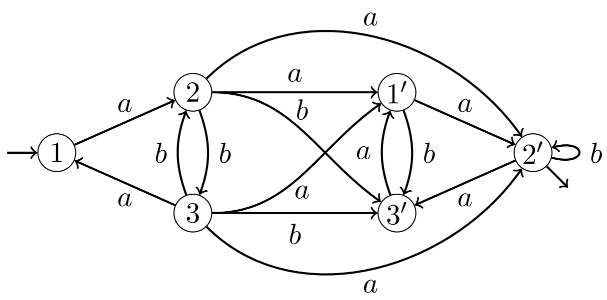


FIGURE 4 – Automate pour $L(\mathcal{A}_1) \bullet L(\mathcal{A}_2)$ (sans ϵ-transitions)

On fait de même pour $(L(\mathcal{A}_2))^2$, on commence par faire une version avec ϵ-transitions :

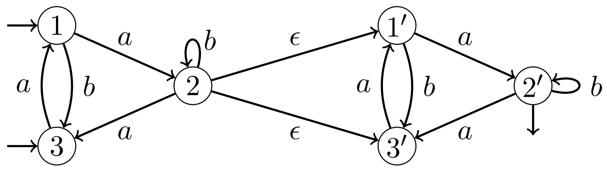


FIGURE 5 – Automate pour $(L(\mathcal{A}_2))^2$ (avec ϵ-transitions)

On élimine ensuite les ϵ-transitions :

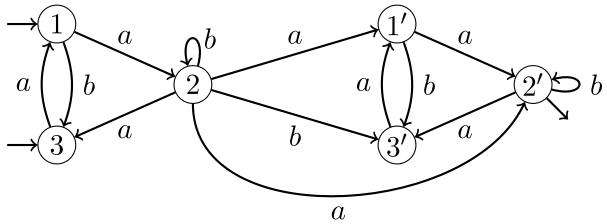


FIGURE 6 – Automate pour $(L(\mathcal{A}_2))^2$ (sans ϵ-transitions)

Pour $(L(\mathcal{A}_2))^*$, on commence par faire une version avec ϵ-transitions :

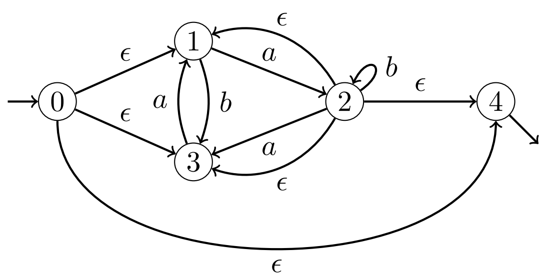


Figure 7 – Automate pour $(L(\mathcal{A}_2))^*$ (avec ϵ-transitions)

On élimine ensuite les ϵ-transitions (et l'état 4 qui devient inaccessible) :

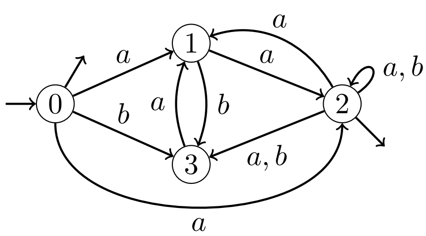


Figure 8 – Automate pour $(L(\mathcal{A}_2))^*$ (sans ϵ-transitions)

Pour $L(\mathcal{A}_1) \setminus L(\mathcal{A}_2)$, on doit construire un automate pour le complémentaire de $\mathcal{A}_2$ puis en faire l'intersection avec $\mathcal{A}_1$. On commence par déterminiser et compléter $\mathcal{A}_2$:

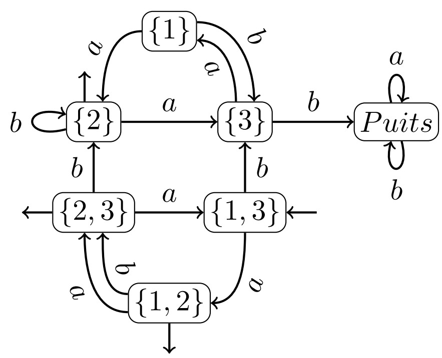


Figure 9 – Automate déterministe complet pour $L(\mathcal{A}_2)$

On en déduit un automate pour $(a+b)^* \setminus L(\mathcal{A}_2)$:

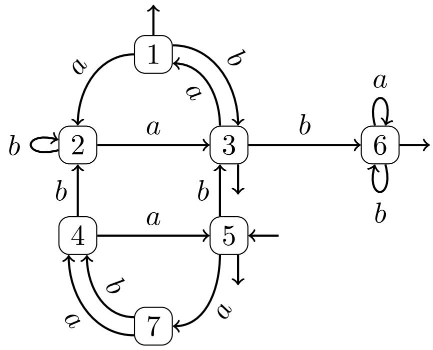


FIGURE 10 – Automate (déterministe complet) pour $(a+b)^* \setminus L(\mathcal{A}_2)$

Il nous reste maintenant à faire l'intersection avec $\mathcal{A}_1$ pour avoir $L(\mathcal{A}_1) \setminus L(\mathcal{A}_2)$

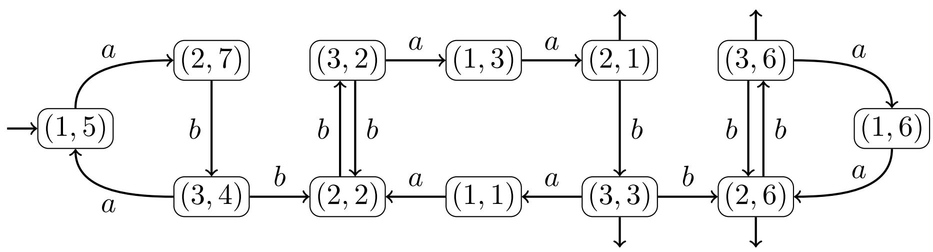


FIGURE 11 – Automate pour $L(\mathcal{A}_1) \setminus L(\mathcal{A}_2)$

Solution de l'exercice 2 :

On considère un automate $\mathcal{A}$ et l'automate $\mathcal{A}'$ obtenu en éliminant les ε-transitions de $\mathcal{A}$. On montre que $L(\mathcal{A}) = L(\mathcal{A}')$.

Pour cela il faut montrer deux directions : si on a un chemin acceptant un mot u dans $\mathcal{A}$, alors on en a un acceptant u dans $\mathcal{A}'$, et vice-versa.

($\Rightarrow$) Supposons qu'il existe un chemin π étiqueté par u dans $\mathcal{A}$ allant d'un état q à un état q'. On raisonne sur le nombre d'ε-transitions dans π. On considère que la dernière transition n'est pas une ε-transition.

Si π ne contient aucune ε-transitions, alors π est par construction un chemin dans $\mathcal{A}'$.

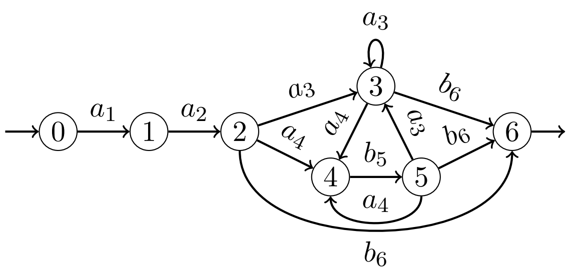
Supposons que π contient i ε-transition et que le résultat est démontré pour tous les chemins de longueur au plus $i-1$. π est de la forme $q \xrightarrow{u_0} q_1 \xrightarrow{\varepsilon}^* q_2 \xrightarrow{a} q_3 \xrightarrow{u_1} q'$ avec u_0 ne terminant pas par une ε-transitions, u_1 ne contenant pas d'ε transition (et donc étant déjà dans $\mathcal{A}'$), et $u_0 a u_1 = u$. Par hypothèse d'induction, il existe un chemin π' de q à q_1 étiqueté par u_0 dans $\mathcal{A}'$. Par construction, il existe une transition $q_1 \xrightarrow{a} q_3$ dans $\mathcal{A}'$. Donc $q \xrightarrow{u_0} q_1 \xrightarrow{a} q_3 \xrightarrow{u_1} q'$ est un chemin étiqueté par u dans $\mathcal{A}'$.

Pour conclure, il reste à considérer les chemin acceptant u terminant par des ε transition. Un tel chemin est de la forme $q_i \xrightarrow{u} q \xrightarrow{\varepsilon}^* q_f$. Par la récurrence précédente, on a un chemin π de q_i à q étiqueté par u dans $\mathcal{A}'$. Par construction, q est final dans $\mathcal{A}'$, et donc u est accepté par π dans $\mathcal{A}'$.

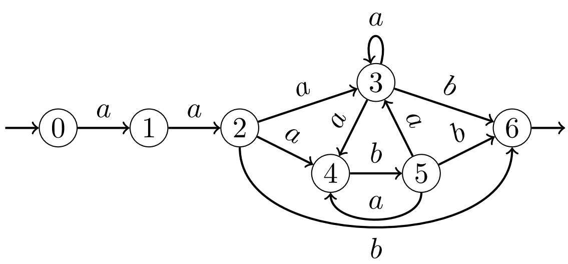
(⇐) Supposons qu'il existe un chemin π étiqueté par u dans $\mathcal{A}'$. Par construction, il existe un chemin étiqueté par des ε-transition d'un état initial de $\mathcal{A}$ vers le premier état de π. Pour chaque transition de π, soit elle est aussi une transition dans $\mathcal{A}$, soit il existe un chemin commençant par la lettre lue, puis constituée d'ε-transition entre les mêmes états. En remplaçant les transitions de π par ces chemins, on obtient un chemin acceptant dans $\mathcal{A}$ et étiqueté par le même mot.

Solution de l'exercice 3 :

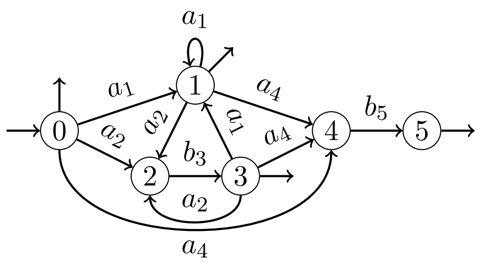
1. $aa(a+ab)^*b$. Après renommage, on obtient : $a_1a_2(a_3+a_4b_5)^*b_6$.



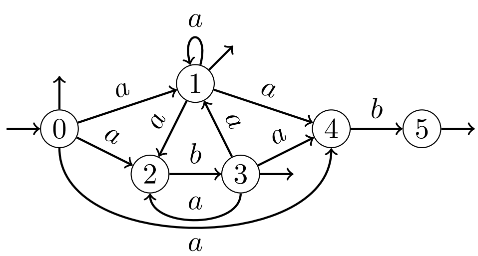
En reprenant l'alphabet initial on obtient :



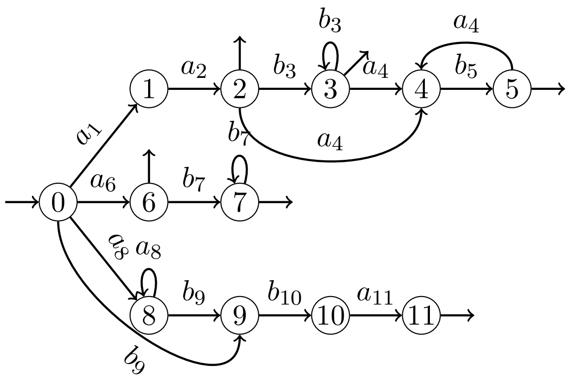
2. $(a+ab)^*(\epsilon+ab)$. Après renommage, on obtient : $(a_1+a_2b_3)^*(\epsilon+a_4b_5)$.



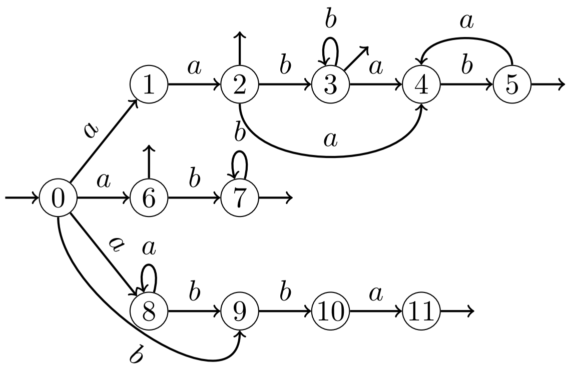
En reprenant l'alphabet initial on obtient :



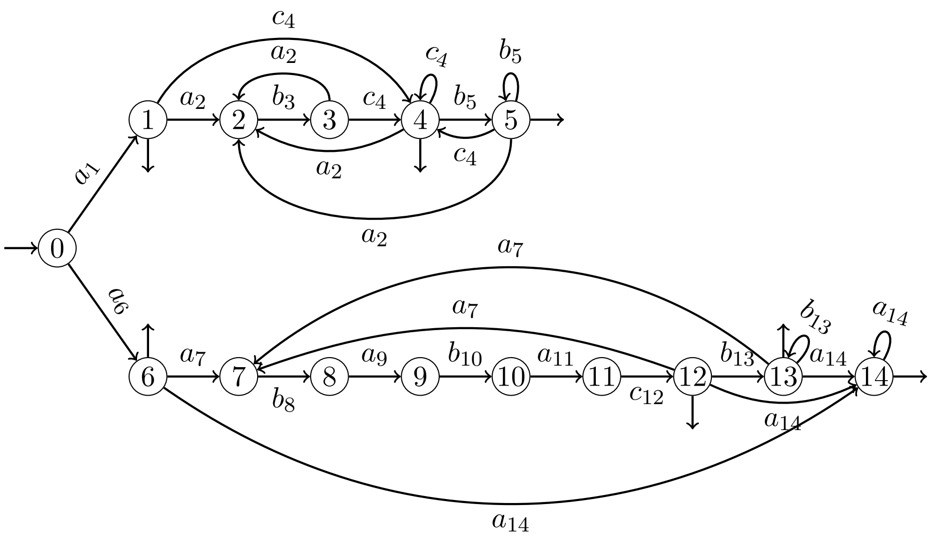
3. $aab^*(ab)^*+ab^*+a^*bba$. Après renommage, on obtient : $a_1a_2b_3^*(a_4b_5)^*+a_6b_7^*+a_8^*b_9b_{10}a_{11}$.



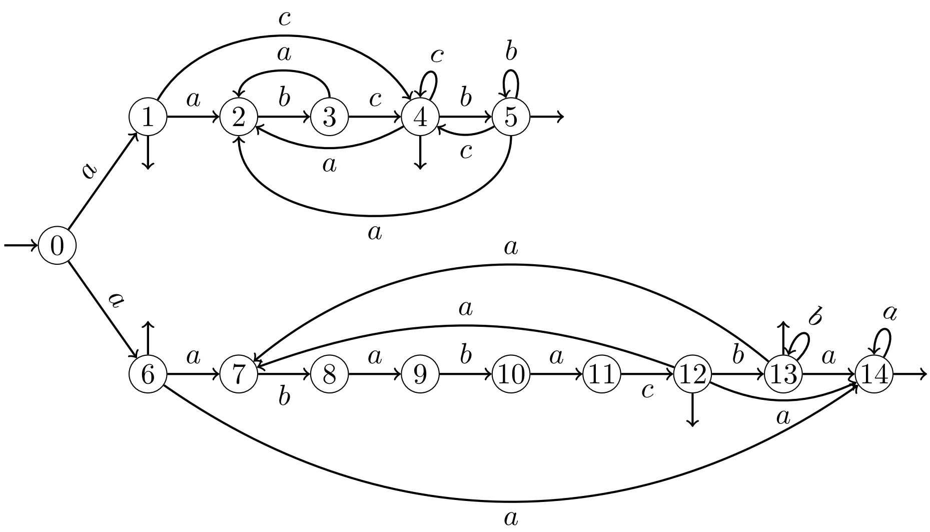
En reprenant l'alphabet initial on obtient :



4. $a((ab)^*cb^*)^* + a(ababacb^*)^*a^*$. Après renommage, on obtient :
$a_1((a_2b_3)^*c_4b_5^*)^* + a_6(a_7b_8a_9b_{10}a_{11}c_{12}b_{13}^*)^*a_{14}^*$.



En reprenant l'alphabet initial on obtient :



Solution de l'exercice 4 :

Équations : Figure 5

On utilise ici la méthode par résolution d'équations avec le lemme d'Arden. Il est important ici de faire une remarque : il existe deux 'versions' symétriques du lemme suivant la forme de l'équation que l'on considère :

Si $X = KX + L$ et $\epsilon \notin K$	Si $X = XK + L$ et $\epsilon \notin K$
alors	*alors*
$X = K^*L$	$X = LK^*$

Suivant la version qu'on utilise la méthode pour générer les équations est légérement différente. On utilise ici la version de droite (on utilisera la version de gauche pour la figure 2)

On considère L_1, L_2, L_3, L_4 les langages définis comme suit : $w \in L_i$ si et seulement si il existe un calcul de l'automate sur w partant de l'état intial et arrivant dans l'état i (attention si on utilisait l'autre version du lemme la définition serait différente). Il découle des transitions qu'on a les équations suivantes :

$$\begin{array}{rcl} L_1 & = & \epsilon \\ L_2 & = & L_1 a \\ L_3 & = & L_2 b + L_3(a+b) \\ L_4 & = & L_1 b + L_2 a + L_4 a \end{array}$$

En substituant L_1 par sa valeur on obtient :

$$\begin{array}{rcl} L_1 & = & \epsilon \\ L_2 & = & a \\ L_3 & = & L_2 b + L_3(a+b) \\ L_4 & = & b + L_2 a + L_4 a \end{array}$$

En substituant L_2 par sa valeur on obtient :

$$\begin{array}{rcl}
L_1 & = & \epsilon \\
L_2 & = & a \\
L_3 & = & ab + L_3(a+b) \\
L_4 & = & b + aa + L_4 a
\end{array}$$

Enfin en appliquant le lemme d'Arden (version de droite) aux deux dernières équations on obtient :

$$\begin{array}{rcl}
L_1 & = & \epsilon \\
L_2 & = & a \\
L_3 & = & ab(a+b)^* \\
L_4 & = & (b+aa)a^*
\end{array}$$

Équations : Figure 6

On utilise cette fois la version de gauche du lemme d'Arden.

On considère L_1, L_2, L_3, L_4 les langages définis comme suit : $w \in L_i$ si et seulement si il existe un calcul de l'automate sur w partant de l'état i et arrivant dans un état final (attention si on utilisait l'autre version du lemme la définition serait différente). Il découle des transitions qu'on a les équations suivantes :

$$\begin{array}{rcl}
L_1 & = & aL_1 + bL_4 + \epsilon \\
L_2 & = & aL_1 + aL_3 + \epsilon \\
L_3 & = & bL_1 \\
L_4 & = & bL_2
\end{array}$$

En substituant L_3 et L_4 on obtient :

$$\begin{array}{rcl}
L_1 & = & aL_1 + bbL_2 + \epsilon \\
L_2 & = & aL_1 + abL_1 + \epsilon \\
L_3 & = & bL_1 \\
L_4 & = & bL_2
\end{array}$$

En substituant L_2 on obtient :

$$\begin{array}{rcl}
L_1 & = & (a + bba + bbab)L_1 + bb + \epsilon \\
L_2 & = & aL_1 + abL_1 + \epsilon \\
L_3 & = & bL_1 \\
L_4 & = & bL_2
\end{array}$$

Enfin en appliquant le lemme d'Arden (version de gauche) à la première équation on obtient : $L_1 = (a + bba + bbab)^*(bb + \epsilon)$.