1 Lemme d'Arden

Exercice 1. Soit $\mathcal{A}$, $\mathcal{B}$ des langages rationnels, montrez que si $\epsilon \notin \mathcal{A}$ alors $\mathcal{X} = \mathcal{A}\mathcal{X} \cup \mathcal{B}$ admet un unique langage $\mathcal{X}$ solution. Exprimez cette solution en fonction de a (resp. b) une expression régulière qui reconnaît $\mathcal{A}$ (resp. $\mathcal{B}$).

Conseil pour aborder cet exercice :

1. *trouvez une première solution et prouvez que c'en est bien une ;*
2. *prouvez que toute solution contient votre première solution ;*
3. *prouvez que votre première solution contient toute solution.*

Solution 1. Nous allons montrer que l'unique solution est $\mathcal{A}^*\mathcal{B} = \{a_1 \dots a_n b \mid (a_i) \in \mathcal{A}^*, b \in \mathcal{B}\}$ ou en forme d'expression régulière a^*b.

$\mathcal{A}^*\mathcal{B}$ est une solution : $\mathcal{A}\mathcal{A}^*\mathcal{B} \cup \mathcal{B} = \mathcal{A}^*\mathcal{B}$.

Montrons par récurrence sur la taille des mots de $\mathcal{X}$ que toute solution est incluse dans a^*b.

Soit $x \in \mathcal{X}$ on a soit $x \in \mathcal{B}$ soit $x \in \mathcal{A}\mathcal{X}$. Si $x \in B$, le résultat est clair. Si $x \in \mathcal{A}\mathcal{X}$ nous avons $a \in \mathcal{A}$ et $x' \in \mathcal{X}$ tels que $x = ax'$. Comme $|a| > 0$, $|x'| < |x|$ et donc par récurrence on a x' dans a^*b donc x aussi.

Enfin toute solution $\mathcal{X}$ contient $\mathcal{B}$ et est stable par concaténation gauche de $\mathcal{A}$ donc a^*b est contenue dans cette solution.

2 Lemme de pompage (ou lemme de l'étoile)

Exercice 2. Les langages suivants sont-ils réguliers ? Justifiez.

1. $\{a^n b^n \mid n \in \mathbb{N}\}$
2. $\{a^m b^n \mid n \equiv m \pmod d\}$ pour un $d \in \mathbb{N}$ donné.
3. $\{a^p \mid p \text{ premier }\}$
4. $\{a^{P(n)} \mid n \in \mathbb{N}\}$ pour $P \in \mathbb{N}[X]$ donné.

Solution 2.

1. Non, d'après le lemme de pompage pour m suffisamment grand on a $a^m b^m$ qui se décomposerait en uvw avec $uv^k w$ dans le langage. Si v panache a et b alors le uv^2w alterne a et b et si w ne contient que a ou que b il y a déséquilibre entre les nombres de a et b.
2. Oui, c'est l'union pour $k = 0, \dots, d-1$ des langages $L_k = (a^k(a^d)^* b^k (b^d)^*)$
3. Non, si pour p suffisamment grand on a a^p implique une décomposition $a^p = uvw$ avec $uv^x w$ dans le langage pour tout x. Si $v = a^d$ on a $a^p + x \times d$ pour tout d or $p + x \times d$ n'est pas premier pour $x = p$ $(p + pd = (d+1)p)$.
4. Par lemme de pompage on a pour n assez grand $a^{P(n)}$ décomposable en uvw avec $|v| > 0$ et $uv^k w$ dans le langage pour tout k. Donc il existe k et d tel que $a^{k \times x + d}$ pour tout x. Ce n'est pas vrai si P est du second degré ou plus (l'écart $P(n+1) - P(n)$ tend vers l'infini). En revanche si P est du premier degré ça marche, pour $P(X) = BX + C$ on a l'expression $(a^B) * a^{(C\%B)}$.

3 Puzzles

Exercice 3. Soient x et y deux mots tels que $xy = yx$, que pouvez-vous dire de la forme de x et y ?

Solution 3. Montrons par récurrence sur $|y|+|x|$ qu'il existe w tel que $x=w^n$ et $y=w^m$ pour $(n,m)\in\mathbb{N}^2$.

Si $|x|=|y|$, on a $x=y$ alors $w=x=y$ avec $n=m=1$ convient.

Sinon on a $xy=yx$ donc on peut trouver z tel que $y=xz$. On a alors $xxz=xzx$ et donc $xz=zx$. Par récurrence on obtient $x=w^n$, $z=w^m$ et donc $y=w^{m+n}$

Exercice 4. Soient $(p,q)\in\mathbb{N}^*$, et $\mathcal{L}$ un langage régulier est-ce que $\frac{p}{q}\mathcal{L}=\left\{u \mid \exists v : uv\in\mathcal{L} \text{ et } |u|=\frac{p}{q}|uv|\right\}$ est nécessairement régulier ?

Solution 4. Oui, nous allons présenter l'automate pour $p\leq q$ et $gcd(p,q)=1$ (les autres cas sont triviaux). L'idée c'est que $|u|=\frac{p}{q}|uv|$ est équivalent à $q|u|=p|uv|$ ou $(q-p)|u|=p|v|$. De la même manière que pour $\frac{1}{2}\mathcal{L}$ où à chaque fois que l'on lit une lettre, on avance d'une lettre en partant de la fin (voir TD de la semaine dernière), ici en lisant p lettres de u on note que l'on avance de $q-p$ lettres dans l'automate en partant de la fin.

Soit $<\Sigma,\mathcal{Q},\delta,i,\mathcal{F}>$ un automate déterministe qui reconnaît $\mathcal{L}$ notre automate est :
$<\Sigma,\mathcal{Q}\times\{0,1,\dots,p-1\}\times 2^{\mathcal{Q}},\eta,(i,0,\mathcal{F}),\{(e,0,f)\mid f\cap\mathcal{F}\neq\emptyset\}>$ Un état (e,l,f) stocke l'état e obtenu par l'automate initial, le nombre l de lettres lues modulo q et les états $g\in f$ tels qu'il existe un mot de taille $\lfloor\frac{p}{q}l\rfloor$ qui mène de g à un état de $\mathcal{F}$.

Les transitions sont donc $\eta((e,p-1,f),c)=(\delta(e,c),0,\{g\mid\exists h\in f, v : |v|=p-q \text{ et } \tilde{\delta}(h,v)=g\})$ et $\eta((e,l,f),c)=(\delta(e,c),l+1,f)$ si $l<q-1$.

4 Lemme d'Arden généralisé

Definition 1. *Un monoïde est une structure algébrique $(E,\bigodot,e)$ telle que :*
(loi interne) Pour $x,y\in E^2$, $x\bigodot y\in E$
(associative) Pour $x,y,z\in E^3$, $(x\bigodot y)\bigodot z=x\bigodot(y\bigodot z)$
(élément neutre) Pour $x\in E$, $x\bigodot e=e\bigodot x=x$

Definition 2. *Un demi-anneau est une structure algébrique $(E,+,\times,0,1)$ telle que :*
- $(E,+,0)$ *est un monoïde commutatif ($a+b=b+a$) ;*
- $(E,\times,1)$ *est un monoïde ;*
- $\times$ *est distributif pour* $+$ *($a\times(b+c)=a\times b+a\times c$ et $(b+c)\times a=b\times a+c\times a$) ;*
- 0 *est absorbant pour* $\times$ *(i.e. $0\times a=a\times 0=0$)*

Exercice 5. Remarquez que $(\mathcal{L}_\Sigma,\cup,/,\emptyset,\epsilon)$ est un demi-anneau où :
- $\mathcal{L}_\Sigma$ est l'ensemble des langages réguliers sur l'alphabet fini Σ ;
- $/$ est la concaténation $X/Y=\{xy\mid x\in X,y\in Y\}$;
- $\cup$ est l'union ;
- $\emptyset$ est le langage vide ;
- ϵ est le langage contenant uniquement le mot vide.

Soit $(M_n(\mathcal{L}_\Sigma),+,\times,0,1)$ l'ensemble des matrices carrés de taille n sur le demi-anneau $(\mathcal{L}_\Sigma,/,\cup,\emptyset,\epsilon)$ avec
- $(M+N)_{i,j}=M_{i,j}\cup N_{i,j}$;
- $(M\times N)_{i,j}=\bigcup_{1\leq k\leq n} M_{i,k}/N_{k,j}$;
- $0_{i,j}=\emptyset$
- $1_{i,j}=\begin{cases}\epsilon & i=j\\ \emptyset & \text{sinon}\end{cases}$
- $M^*=\sum_{i\in\mathbb{N}}M^i$

Remarquez que $(M_n(\mathcal{L}_\Sigma),+,\times,0,1)$ est aussi un demi-anneau.

Note : Tout demi-anneau peut se transformer de la même manière en un demi-anneau de matrices.

Exercice 6. Soit $\mathcal{A} \in M_n(\mathcal{L}_\Sigma)$ avec $\forall i, j\ \epsilon \notin A_{i,j}$ et $\mathcal{B} \in M_{n,1}(\mathcal{L}_\Sigma)$ un vecteur de taille n montrez qu'il n'existe qu'une unique solution $\mathcal{X} \in M_{n,1}(\mathcal{L}_\Sigma)$ à $\mathcal{X} = \mathcal{A}\mathcal{X} \cup \mathcal{B}$ (i.e. $\forall i : \mathcal{X}_i = \left(\bigcup_j \mathcal{A}_{i,j}/\mathcal{X}_j\right) \cup \mathcal{B}_i$).

Solution 6.

L'unique solution est $\mathcal{A}^*\mathcal{B} = \bigcup_n \mathcal{A}^n\mathcal{B}$. En effet, c'est une solution, et toutes les solutions doivent contenir $\mathcal{B}$ et être stables par concaténation gauche de $\mathcal{A}$ donc toutes les solutions contiennent $\mathcal{A}^*\mathcal{B}$ et par conséquence pour tout j, une solution $\mathcal{X}$ est telle que $(\mathcal{A}^*\mathcal{B})_j \subseteq \mathcal{X}_j$.

Maintenant soit $\mathcal{X}$ une solution avec $\mathcal{X} \neq \mathcal{A}^*\mathcal{B}$, soient j et $(x_1, \dots, x_n) \in \mathcal{X}$ tels que $|x_j|$ est minimal et $x_j \in \mathcal{X}_j \setminus (\mathcal{A}^*\mathcal{B})_j$. $\mathcal{X} = \mathcal{A}\mathcal{X} \cup \mathcal{B}$ donc comme par définition $\mathcal{B}_j \subseteq \mathcal{X}_j$ il existe $a_{j,i}x_i$ et $x_i \in \mathcal{X}_i$. Comme $|x_j|$ est minimal et $|a_{i,j}| > 0$ on a $x_i \in \mathcal{A}^*\mathcal{B}$ mais alors $x_j \in (\mathcal{A}^*\mathcal{B})_j$ ce qui est contradictoire.

Exercice 7. *(Pivot de Gauss)* Étant donné $\mathcal{A} \in M_n(\mathcal{L}_\Sigma)$ et $\mathcal{B} \in M_{n,1}(\mathcal{L}_\Sigma)$ montrez que si $n > 1$ les langages $(\mathcal{A}^*\mathcal{B})_i$ pour $i \in \{1, \dots, n-1\}$ peuvent s'exprimer sous la forme $\mathcal{A}'^*\mathcal{B}'$ pour $\mathcal{A}' \in M_n(\mathcal{L}_\Sigma), \mathcal{B}' \in M_{n,1}(\mathcal{L}_\Sigma)$.

Solution 7. On a $\mathcal{X}_n = \bigcup_i \mathcal{A}_{n,i}\mathcal{X}_i \cup \mathcal{B}_n$ donc $\mathcal{X}_n = \bigcup_j \mathcal{A}_{n,i}\mathcal{X}_i\,(\mathcal{A}_{n,n}\mathcal{X}_n \cup \mathcal{B}_n)$ donc $\mathcal{X}_n = \mathcal{A}^*_{n,n}\left(\mathcal{B}_n \bigcup_j \mathcal{A}_{n,i}\mathcal{X}_i\right)$.

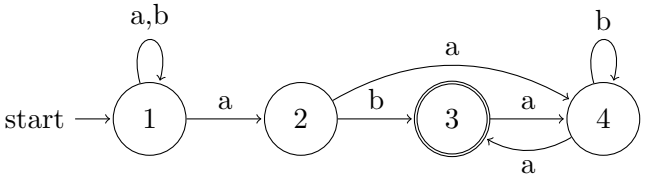
On avait $\mathcal{X}_i = \bigcup_j \mathcal{A}_{i,j}\mathcal{X}_j$ on a $\mathcal{X}_i = \bigcup_{j<n} \mathcal{A}_{i,j}\mathcal{X}_j \cup \mathcal{A}_{i,n}(\mathcal{A}^*_{n,n}\mathcal{B}_n \cup \bigcup \mathcal{A}^*_{n,n}A_{n,j}X_j) \cup \mathcal{B}_i$ donc nous avons $\mathcal{X}_i = \bigcup_{j<n}(\mathcal{A}_{i,j} \cup \mathcal{A}_{i,n}\mathcal{A}^*_{n,n}\mathcal{A}_{n,j})\mathcal{X}_j \cup (\mathcal{B}_i \cup A_{i,n}A^*_{n,n}\mathcal{B}_n)$.

En posant $\mathcal{A}'_{i,j} = \mathcal{A}_{i,j} \cup \mathcal{A}_{i,n}\mathcal{A}^*_{n,n}\mathcal{A}_{n,j}$ et $\mathcal{B}'_i = B_i \cup A_{i,n}A^*_{n,n}\mathcal{B}_n$ nous avons bien le résultat escompté.

Exercice 8. En déduire un algorithme permettant de retrouver une expression régulière à partir d'un automate.

Solution 8. Un automate $< \Sigma, \mathcal{Q}, \delta, i, \mathcal{F} >$ peut être vu comme un système d'équations avec $A_{i,j} = c_1| \dots |c_k$ où les c_l sont tels que $q_j \in \delta(q_i, c_l)$ et $B_i = \begin{cases} \epsilon & q_i \in \mathcal{F} \\ \emptyset & \text{sinon} \end{cases}$. À partir de là on applique simplement le pivot de Gauss jusqu'à obtenir une seule équation correspond à l'état initial $\mathcal{X} = \mathcal{A}\mathcal{X} \cup \mathcal{B}$ et la solution est $\mathcal{A}^*\mathcal{B}$.

Exercice 9. Donner l'expression régulière pour l'automate suivant :



Solution 9.

$(a\|b)$	a	$\emptyset$	$\emptyset$	$\emptyset$
$\emptyset$	$\emptyset$	b	a	$\emptyset$
$\emptyset$	$\emptyset$	$\emptyset$	a	ϵ
$\emptyset$	$\emptyset$	a	b	$\emptyset$

$(a\|b)$	a	$\emptyset$	$\emptyset$
$\emptyset$	$\emptyset$	$b\|(ab^*a)$	$\emptyset$
$\emptyset$	$\emptyset$	ab^*a	ϵ

$(a\|b)$	a	$\emptyset$
$\emptyset$	$\emptyset$	$(b\|(ab^*a))(ab^*a)^*$

$(a\|b)$	$a(b\|(ab^*a))(ab^*a)^*$

Le langage est donc $(a|b)^*a(b|(ab^*a))(ab^*a)^*$

5 Monoïde d'un langage

Definition 3. *Étant donné un alphabet ambiant Σ et un langage $\mathcal{L}$ (pas nécessairement régulier) on note $\mu_\mathcal{L}$ le morphisme syntaxique de $\mathcal{L}$ définit ainsi : $\mu_\mathcal{L}(u) = \{(x, y) \in \Sigma^{*2} \mid xuy \in \mathcal{L}\}$.*

Exercice 10. Montrer que $\mu_\mathcal{L}$ est compatible avec la concaténation c'est à dire $\forall u, v, w : \mu_\mathcal{L}(v) = \mu_\mathcal{L}(w) \Rightarrow (\mu_\mathcal{L}(vu) = \mu_\mathcal{L}(wu)) \wedge (\mu_\mathcal{L}(uv) = \mu_\mathcal{L}(uw))$.

Solution 10. Soient u, v, w avec $\mu_{\mathcal{L}}(w) = \mu_{\mathcal{L}}(v)$ on a $(x, y) \in \mu_{\mathcal{L}}(vu)$ équivalent $xvuy \in \mathcal{L}$ donc $(x, uy) \in \mu_{\mathcal{L}}(v) = \mu_{\mathcal{L}}(w)$ donc $(x, y) \in ml(wu)$ et donc $\mu_{\mathcal{L}}(vu) \subseteq \mu_{\mathcal{L}}(wu)$. Tous les autres cas sont parfaitement symétriques donc on obtient le résultat voulu.

Exercice 11. Pour chaque élément $e \in \mu_{\mathcal{L}}(\Sigma^*)$ on choisit un représentant $r(e)$ tel que $\mu_{\mathcal{L}}(r(e)) = e$, et on définit $u \odot v = \mu_{\mathcal{L}}(r(u)) \odot \mu_{\mathcal{L}}(r(v)) = \mu_{\mathcal{L}}(r(u)r(v))$.

Justifier que :

— $\odot$ ne dépend pas du choix de r.

— la structure $(\mu_{\mathcal{L}}(\Sigma^*), \odot, \mu_{\mathcal{L}}(\epsilon))$ induite par $\mu_{\mathcal{L}}$ est bien un monoïde

et donc $\mu_{\mathcal{L}}$ est bien un morphisme entre le monoïde libre sur Σ et le monoïde induit par $\mu_{\mathcal{L}}$.

Note : tout monoïde $(M, \odot, e)$ compatible avec une fonction μ induit une structure de monoïde sur $\mu(M)$.

Solution 11.

— Soient r et r', d'après la question précédente on a : $u \odot_r v = \mu(r(u)) \odot_r \mu(r(v)) = \mu(r(u)r(v)) = \mu(r'(u)r(v)) = \mu(r'(u)r'(v)) = u \odot_{r'} v$

— $\odot$ est bien une loi interne, montrons son associativité : $(x \odot y) \odot z = (r(x) \odot r(y)) \odot r(z) = \mu_{\mathcal{L}}(r(\mu_{\mathcal{L}}(xy))r(z))$ mais $\mu_{\mathcal{L}}(r(\mu_{\mathcal{L}}(xy))z) = \mu_{\mathcal{L}}(xyz)$ car $\mu_{\mathcal{L}}(xy) = \mu_{\mathcal{L}}(r(\mu_{\mathcal{L}}(xy))$. De la même manière on a $x \odot (y \odot z) = \mu_{\mathcal{L}}(xyz)$. Enfin $\mu_{\mathcal{L}}(r(x)r(\mu_{\mathcal{L}}(\epsilon))) = \mu_{\mathcal{L}}(r(x)) = \mu_{\mathcal{L}}(r(\mu_{\mathcal{L}}(\epsilon))r(x))$ donc $\mu_{\mathcal{L}}(\epsilon)$ est bien associatif.

Exercice 12. Montrer que $\mu_{\mathcal{L}}(\Sigma^*)$ est fini quand $\mathcal{L}$ est régulier.

Solution 12. Si $\mathcal{L}$ est régulier, il existe un automate déterministe fini complet $< \Sigma, \mathcal{Q}, \delta, i, \mathcal{F} >$ qui reconnaît $\mathcal{L}$. Pour un u donné on pose $t(u) = \{(p, \tilde{\delta}(p, u)) \in \mathcal{Q}^2 \mid q \in \tilde{\delta}(p, u)\}$ montrons que $t(u) = t(v)$ implique $\mu_{\mathcal{L}}(u) = \mu_{\mathcal{L}}(v)$. Par symétrie, montrons juste que $\mu_{\mathcal{L}}(u) \subseteq \mu_{\mathcal{L}}(v)$.

Soient $u, v, x, y \in \Sigma^{*2}$ et $(x, y) \in \mu_{\mathcal{L}}(u)$ on a $xuy \in Lc$ donc $\tilde{\delta}(i, xuy) \in \mathcal{F}$ mais $\tilde{\delta}(i, xuy) = \tilde{\delta}(q_x, uy)$ avec $q_x = \tilde{\delta}(i, x)$. Donc $\tilde{\delta}(i, xuy) = \tilde{\delta}(\tilde{\delta}(q_x, u), y) = \tilde{\delta}(\tilde{\delta}(q_x, v), y) = \tilde{\delta}(i, xvy)$ et donc $xvy \in \mathcal{L}$.

Exercice 13. Montrer que si $\mu_{\mathcal{L}}(\Sigma^*)$ est fini alors $\mathcal{L}$ est régulier.

Solution 13. Si $\mu_{\mathcal{L}}(\Sigma^*) = \{q_1, \ldots, q_n\}$ on pose $i = \mu_{\mathcal{L}}(\epsilon)$, $\delta(q_i, c) = q_i \odot \mu_{\mathcal{L}}(c)$ et $\mathcal{F} = \{q \in \mu_{\mathcal{L}}(\Sigma^*) \mid (\epsilon, \epsilon) \in q\}$ et alors $< \Sigma, \mu_{\mathcal{L}}(\Sigma^*), \delta, i, \mathcal{F} >$ est un automate qui reconnaît $\mathcal{L}$.

Exercice 14. À quoi ressemble le monoïde syntaxique des mots bien parenthèsés ?

Solution 14. À $\mathbb{Z}^2$: un mot est envoyé sur $(x, y) \in \mathbb{Z}^2$ quand il ferme x parenthèses ouvertes avant lui et laisse ouvertes y parenthèses et on a la loi $(x, y) \odot (w, z) = (x + max(w - y, 0), z + max(y - w, 0))$.