Automates avancés et applications

Master 1

Examen du 16 mai 2018 — corrigé incomplet.

1 Applications des cours

1.1 Automates à pile

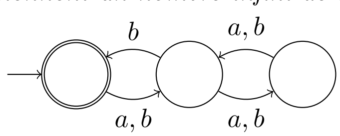
Trouvez un automate à pile qui reconnait le langage $U = \{a^i b^j c^k | i = j \vee j = k\}$ *vu en partiel. Précisez s'il le reconnait par pile vide ou état final.* L'idée est de faire un automate qui choisit par ε-transition entre 2 branches, une pour $\{a^i b^i c^*\}$, autre pour $\{a^* b^j c^j\}$¿ la première branche empile le symbole 1 en lisant les a, dépile en lisant les b, lit une séquence de c sans regarder la pile, et reconnait en enlevant Z_0 du fond de pile. L'autre branche est similaire.

Faire un dessin.

1.2 Mots infinis

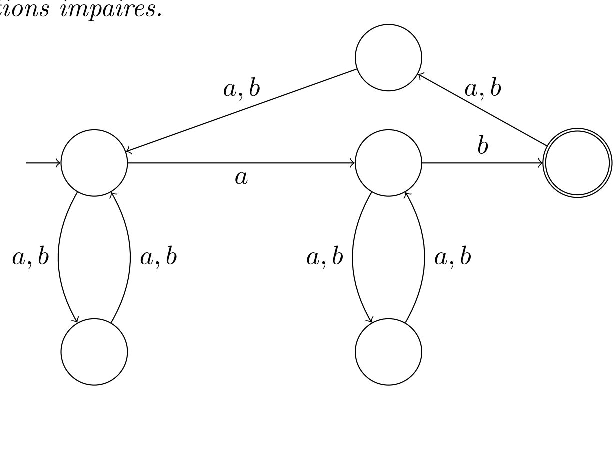
On suppose que la 1ère lettre du mot a le numéro 1. *Trouvez un automate de Büchi et une expression ω-régulière pour les ω-langages sur $\{a, b\}$ suivants :*

— *tous les mots qui contiennent un nombre infini de b sur des positions paires.*



$\big((a + b)(a + b)^{2*} b\big)^\omega$

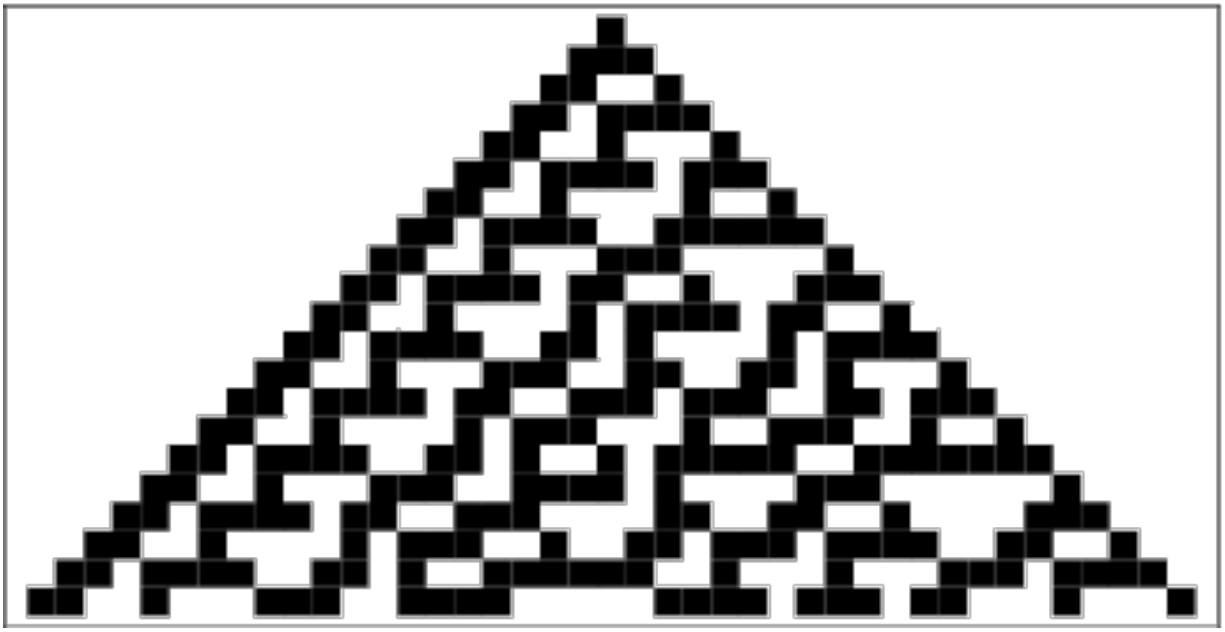
— *tous les mots qui contiennent un nombre infini de b sur des positions paires et un nombre infini de a sur des positions impaires.*



$\big((a + b)^{2*} a (a + b)^{2*} b\big)^\omega$

1.3 Automates cellulaires

Trouvez le numéro de Wolfram de l'automate cellulaire à partir du diagramme spatio-temporel :



on observe que

$$f(111) = 0; f(110) = 0; f(101) = 0; f(100) = 1; f(011) = 1; f(010) = 1; f(001) = 1; f(000) = 0;$$

et on déduit le numéro de Wolfram : 000111110 en base 2 ; c-à-d 30.

Possède-t-il des jumeaux ? Un Jardin d'Eden ? Justifiez. Les jumeaux sont deux configurations finies avec le même successeurs, par exemple 01110 et 01111 dont le successeur est 11000. Par théorème classique cela implique l'existence d'un jardin d'Eden.

2 Des petits raisonnements

Soit le langage L de tous les mots sur $\{1, 2, 3\}$ avec le même nombre de chaque chiffre.

— *Démontrez qu'il n'est pas hors contexte.* Preuve banale par pumping lemma.

— *Donnez une grammaire qui l'engendre.* Par point précédent, cette grammaire ne peut pas être hors contexte.
On va engendrer $(ABC)^*$, permuter les lettres arbitrairement et transformer en terminaux :

$$S \to SABC|\varepsilon; AB \leftrightarrow BA; AC \leftrightarrow CA; BC \leftrightarrow CB; A \to 1; B \to 2; C \to 3.$$

2.1 On utilise les automates à pile

Démontrez que l'intersection d'un langage hors contexte L_1 et un langage régulier L_2 est hors contexte. L'idée est simple : on prend un automate à pile A_1 qui reconnait L_1 (par état final) et un automate fini A_2 qui reconnait L_2 et on construit leur automate produit (il aura une pile) qui acceptera un mot si simultanément A_1 et A_2 arriveront en état final sur ce mot.

La construction détaillée...

2.2 Mots périodiques et leurs applications

1. *Démontrer que chaque langage ω-régulier non-vide contient un mot ultimement périodique (c-à-d de la forme uv^ω).*

Soit L un langage ω-régulier non-vide accepté par un automate de Büchi $\mathcal{A}$. Comme L est non-vide, il contient un mot infini x. Un calcul accepteur de $\mathcal{A}$ sur x doit exister. Ce calcul commence dans l'état initial i de $\mathcal{A}$ et visite infiniment souvent l'ensemble F de ses états finals. Comme F est fini, ce calcul doit visiter un certain $f \in F$ au moins deux fois. On a donc vu que le début de calcul sur x a nécessairement la forme

$$i \xrightarrow{u} f \xrightarrow{v} f \rightarrow \cdots$$

avec u, v deux sous-mots finis de x. En particulier on déduit que $i \xrightarrow{u} f$ and $f \xrightarrow{v} f$ pour i initial et f final.

Maintenant on construit un nouveau mot infini $y = uv^\omega$. Par construction il est ultimement périodique. Il possède un calcul accepteur suivant :

$$i \xrightarrow{u} f \xrightarrow{v} f \xrightarrow{v} f \xrightarrow{v} f \xrightarrow{v} f \xrightarrow{v} f \xrightarrow{v} f \xrightarrow{v} f \xrightarrow{v} \cdots$$

On déduit que y est accepté par $\mathcal{A}$ et appartient donc à L.

2. **Lemme 1** *Soit $\mathcal{A}$ un automate de Büchi avec n états acceptant au moins un mot infini. Alors il existe deux mots finis u et v, de taille inférieure à $n+1$, tels que uv^ω est accepté par $\mathcal{A}$.*

   Dans le point précédents on a trouvé deux mots finis u, v tels que $i \xrightarrow{u} f$ and $f \xrightarrow{v} f$ pour i initial et f final. Il existe donc un chemin de i à f. Par conséquent, il existe un chemin simple (qui ne passe jamais par le même état). Un tel chemin contient au plus n transitions. Soit u' le mot qui étiquette ce chemin. On a

   $$|u'| \leq n; \quad i \xrightarrow{u'} f$$

   De la même manière on trouve une boucle simple de f vers f, étiquetée par un mot v' tel que

   $$|v'| \leq n; \quad f \xrightarrow{v'} f$$

   Le même raisonnement que dans le point précédent montre que le mot infini $y' = u'v'^\omega$ est accepté par $\mathcal{A}$.

3. *Donner un exemple de langage ω-régulier de mots infinis, qui ne peut être reconnu par aucun automate avec moins de 1000 états.*

   Exemple : $L = (ba^{1000})^\omega$. Ce langage est bien sûr ω-régulier et contient un seul mot infini $z = (ba^{1000})^\omega$.

   Preuve : Supposons que L est accepté par un automate avec < 1000 états. Par le lemme précédent L contient un mot de la forme uv^ω avec $|u|, |v| < 1000$. Or z est le seul mot dans L. Par conséquent il faut que $z = uv^\omega$

   On déduit donc que le mot infini $z = (ba^{1000})^\omega$ admet une représentation $z = uv^\omega$ avec $|u|, |v| < 1000$. C'est impossible, parce que la plus courte période de z est 1001. **Comment le démontrer rigoureusement ?**

4. *Donner un exemple de langage non-ω-régulier de mots infinis* **Exemple :** L contient un seul mot : la séquence des décimales de π : 314159265359...

   Preuve : L est non-vide et son unique élément n'est pas ultimement périodique. Or si L était ω-régulier, il devrait contenir un mot ultimement périodique. Donc L n'est pas ω-régulier.

3 Toute une théorie : mots bi-infinis

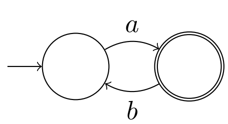
Soit $\mathcal{A} = (Q, \Sigma, I, \Delta, F)$ un automate fini (I, F sont deux ensembles d'états). Un calcul sur un mot bi-infini $\ldots a_{-1}a_0a_1 \ldots$ est une suite bi-infinie

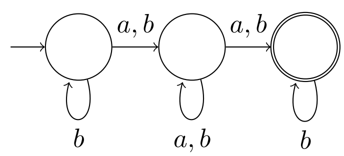
$$\cdots q_{-1} \xrightarrow{a_{-1}} q_0 \xrightarrow{a_0} q_1 \xrightarrow{a_1} q_2 \cdots$$

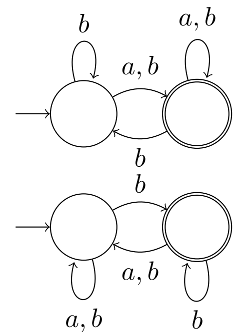
telle que : $\forall i, (q_i, a_i, q_{i+1}) \in \Delta$. Un calcul est *accepteur* s'il existe une infinité de $i \le 0$ tels que $q_i \in I$ et une infinité de $i \ge 0$ tels que $q_i \in F$. Un mot bi-infini est *reconnu* par l'automate $\mathcal{A}$ s'il possède un calcul accepteur. On définit de façon immédiate les langages *reconnaissables* de mots bi-infinis sur un alphabet Σ.

1. *Donnez les automates sur l'alphabet $\{a, b\}$ qui reconnaissent les langages de mots bi-infinis suivants :*
   (a) *Le seul mot $\ldots abababababab \ldots$*
   (b) tous les mots avec un nombre fini de a ;
   (c) tous les mots avec un nombre infini de b.

   Les deux premiers cas sont faciles. Pour le troisième il y a une subtilité : on peut avoir une infinité de b vers la gauche, vers la droite ou des deux cotés.







2. *Soit L_0 l'ensemble de tous les mots bi-infinis sur $\{a, b\}$ avec la lettre a en position 0. Est-il reconnaissable ?*

   Si on décale un calcul accepteur de k positions, il reste un calcul accepteur. Le mot accepté est ainsi décalé de k positions. On déduit le lemme suivant :

   Lemme 2 *Un langage bi-infini reconnaissable est invariant par décalage ; avec chaque mot w il contient le mot $v = dec(v)$ tel que $\forall i(v_i = w_{i+k})$.*

   Or le langage de l'énoncé n'est pas invariant par décalage : il contient un seul mot, mais pas ses décalés. Donc il n'est pas reconnaissable.

3. *Montrez que les langages bi-infinis reconnaissables sont clos par union.* Il suffit de faire une union disjointe des automates.

4. *Les expressions régulières bi-infinies (BRE) et leur sémantique.* Soit RE les expression rationnelles usuelles sans ε sur un alphabet Σ (pour les langages réguliers de mots finis), on rappelle la définition de la syntaxe (mais on omet la sémantique)

$$RE ::== a | RE + RE | RE \cdot RE | RE^*,$$

   où $a \in \Sigma$. On peut maintenant définir les BRE

$$BRE ::== BRE + BRE | RE^{-\omega} \cdot RE \cdot RE^{\omega} | RE^{-\omega} \cdot RE^{\omega}$$

   La sémantique des BRE est définie par induction structurelle avec deux cas.
   — l'union est facile $[e + f] = [e] \cup [f]$ pour BRE e et f.
   — l'autre opération et un peu plus subtile. Soient e, f, g trois RE. Alors le langage

$$[e^{-\omega} \cdot f \cdot g^{\omega}]$$

   consiste de tous les mots bi-infinis w tels qu'il existe une séquence d'indices bi-infinie croissante i_n avec trois propriétés
   — pour tout $n < 0$ le sous-mot de w de la position i_n jusqu'à $i_{n+1} - 1$ appartient à $[e]$;
   — (pour $n = 0$) le sous-mot de w de la position i_0 jusqu'à $i_1 - 1$ appartient à $[f]$;

— pour tout $n > 0$ le sous-mot de w de la position i_n jusqu'à $i_{n+1} - 1$ appartient à $[g]$. (On remarque que cette définition est bien invariante par décalage)

— le dernier cas est similaire au précédent (mais plus simple).

Les BRE pour les langages du premier point de cet exercice.

$$(ab)^{-\omega} \cdot (ab)^{\omega}$$
$$b^{-\omega} \cdot (a+b)^* \cdot b^{\omega}$$
$$(b(a+b)^*)^{-\omega}(a+b)^{\omega} + (a+b)^{-\omega}(b(a+b)^*)^{\omega}$$

5. *Un langage de mots bi-infinis est reconnaissable si et seulement si il peut être exprimé par une BRE.*

Reconnaissable $\Rightarrow$ rationnel Dans un automate A, pour deux états p et q soit L_{pq} l'ensemble de tous les mots finis non-vides qui mènent de p à q. Par théorème de Kleene ce langage peut être défini par une RE.

Le langage de mots bi-infinis reconnus par A est :

$$\bigcup_{i \in I, f \in F} {L_{ii}}^{-\omega} \cdot L_{if} \cdot {L_{ff}}^{\omega}.$$

En remplaçant chaque L_{pq} par son RE et le $\bigcup$ par une somme on obtient une BRE pour L.

Rationnel $\Rightarrow$ reconnaissable Induction structurelle. Il y a deux cas

— l'union est facile l'automate pour BRE $e+f$ est l'union disjointe des automates pour e et pour f.

— étant donnée une BRE $e^{-\omega} \cdot f \cdot g^{\omega}$ on construit d'abord les automates finis A, B, C pour les RE e, f, g. On les fait "normaux", c-à-d avec un seul état initial sans transitions entrantes, et avec un seul état final sans transition sortantes. On fait l'union disjoint de A, B et C, et puis on identifie certains états :

— i_A avec f_A ;

— f_A avec i_B ;

— f_B avec i_C ;

— i_C avec f_C.

Le seul état initial de l'automate obtenu et i_A, le seul état final est f_C.