Langages formels, calculabilité et complexité
TD6

7 novembre 2014

Exercice 1 Lemme d'Ogden (*avancé*)

L'objectif de cet exercice est de démontrer le résultat suivant :

Lemme d'Ogden. Soit L un langage algébrique. Il existe un entier N tel que pour tout mot $z \in L$ dans lequel au moins N positions distinctes sont distinguées, il est possible de décomposer z sous la forme $z = uxvyw$ avec

- x ou y contient au moins une position distinguée,
- xvy contient au plus N positions distinguées,
- pour tout $i \geq 0$, $ux^i vy^i w \in L$.

Soit T un sous-arbre d'un arbre de dérivation selon une grammaire CNF. On suppose distinguée certaines feuilles de T. On appelle *spécial* un nœud de T ayant deux fils, tel que chacun de ses fils contienne au moins une feuille distinguée.

1. Soit T un sous-arbre d'un arbre de dérivation selon une grammaire CNF. On suppose qu'au moins 2^h feuilles distinctes de T sont distinguées.
   Montrer qu'il existe un chemin, d'une feuille à la racine, passant par au moins h nœuds spéciaux et tel que pour tout i, le i-ème nœud spécial ait au plus 2^i descendants distingués.
2. On considère une grammaire CNF engendrant le langage L. Montrer qu'il existe un entier N tel que :
   – pour tout mot $w \in L$ dans lequel au moins N positions sont distinguées,
   – pour tout arbre de dérivation de w,
   il existe deux nœuds spéciaux b_1 et b_2 tels que
   1. b_1 est un ancêtre de b_2,
   2. b_1 est un ancêtre d'au plus N feuilles distinguées,
   3. b_1 et b_2 sont étiquetés par la même variable.
3. En déduire le lemme d'Ogden.

On s'intéresse au langage $L = \{a^i b^j c^k d^l \mid i = 0 \text{ ou } j = k = l\}$.

4. Montrer que pour tout $N \geq 1$ et tout mot $z \in L$ avec $|z| \geq N$, il existe une décomposition $z = uxvyw$ telle que
   – $|xy| \geq 1$
   – $|xvy| \leq N$
   – pour tout $i \geq 0$, $ux^i vy^i w \in L$.
   En d'autres termes, L satisfait les hypothèses du lemme de l'étoile simple.
5. Montrer que L n'est pas algébrique.

Exercice 2 Langage inhéremment ambigu (*avancé*)

Soit $L = \{a^m b^m c^n \mid m, n \geq 0\} \cup \{a^m b^n c^n \mid m, n \geq 0\}$. On veut montrer qu'il n'existe pas de grammaire non-ambiguë qui engendre ce langage.

1. Donner une grammaire engendrant ce langage.
2. Déduire du lemme d'Ogden appliqué à $z = a^k b^k c^{k+k!}$ pour un k assez grand la forme des dérivations possible pour z.
3. En déduire que le mot $a^{k+k!} b^{k+k!} c^{k+k!}$ se dérive d'au moins deux manières différentes.

Exercice 3 Langages algébriques (*base*)

Les langages suivants sur l'alphabet $\{a,b,c\}$ sont-ils algébriques ?

1. $\{a^n b^m \mid n \geq m\}$
2. $\{a^p \mid p \text{ est premier}\}$
3. $\{w \in \{a,b\}^* \mid |w|_a/|w|_b = 3/2\}$
4. $\{wcw^R \mid w \in \{a,b\}^*\}$ où w^R dénote le miroir de w
5. $\{vcw \mid v,w \in \{a,b\}^* \text{ et } w \neq v^R\}$
6. $\{wcw \mid w \in \{a,b\}^*\}$

Exercice 4 Automates de Büchi semi-déterministes (*avancé*)

Un automate $\mathscr{A} = (Q,A,\delta,I,F)$ est semi-déterministe s'il existe une partition $Q = N \cup D$ de l'ensemble de ses états telle que tout état final est dans D et l'automate réstreint $(D,A,\delta|_{D\times A\times D},\{d\},F)$ est déterministe pour tout $d \in D$. Le but de cet exercice est de montrer que les automates de Büchi semi-déterministes sont aussi expressifs que les automates non-déterministes.

Pour un automate $\mathscr{A} = (Q,A,\delta,I,F)$, on définit l'automate semi-déterministe $\mathscr{B} = (N \cup D, A, \delta_B, \{I\}, F_B)$ où $N = \mathscr{P}(Q)$, $D = \mathscr{P}(Q) \times \mathscr{P}(Q)$, $F_B = \{(S,S) \mid S \in \mathscr{P}(Q), S \neq \emptyset\}$ et δ_B est défini comme suit :

- $\delta_B(S,a) = \delta(S,a) \cup \big\{(\{q\},\emptyset) \mid q \in \delta(S,a)\big\}$ pour tout $S \in N$ et $a \in A$
- $\delta_B((S,T),a) = \big\{\big(\delta(S,a), \delta(T,a) \cup (\delta(S,a) \cap F)\big)\big\}$ pour tout $(S,T) \in D$ avec $S \neq T$ et $a \in A$
- $\delta_B((S,S),a) = \big\{\big(\delta(S,a), \delta(S,a) \cap F\big)\big\}$ pour tout $(S,S) \in D$ et $a \in A$

1. Montrer que tout mot reconnu par $\mathscr{B}$ est reconnu par $\mathscr{A}$, au sens de Büchi.
2. Montrer que tout mot reconnu par $\mathscr{A}$ est reconnu par $\mathscr{B}$.
3. En déduire qu'un langage L est reconnu par un automate de Büchi si et seulement s'il existe un nombre fini de langages rationnels $K_1, L_1, \ldots, K_r, L_r$ tel que

$$L = \bigcup_{s=1}^{r} K_s \cdot \overrightarrow{L_s} \ .$$