Langages Formels

Licence Informatique/Math-Info – ENS Paris-Saclay

Examen du 30 mai 2024

Toutes les réponses devront être correctement justifiées. Temps : 2h. Le chiffre à côté d'une question indique sa difficulté ou longueur.

Rappels

Une grammaire algébrique G est donné par $\langle \Sigma, V, P, S \rangle$, pour un alphabet de *terminaux* Σ, un ensemble de variables V, des productions $P \subseteq V \times (V \cup \Sigma)^*$, tous fini, et $S \in V$ le symbole de départ.

Un automate à pile (AAP) $\mathcal{A}$ est donné par $\langle Q, \Sigma, Z, T, q_0 z_0, F \rangle$, avec Q les états, Σ l'alphabet, Z les symboles de pile, $T \subseteq QZ \times (\Sigma \cup \{\varepsilon\}) \times QZ^*$ les transitions (tous fini), $q_0 z_0$ la configuration initiale, et $F \subseteq Q$.

1 Jeu solitaire

On considère l'alphabet $\Sigma = \{a, b, c\}$ et un mot w. Un *bloc* de w est un facteur maximal de longueur au moins 3 composé d'une seule lettre. Un tour du jeu consiste a supprimer un tel bloc. Le but du jeu est d'atteindre le mot vide, et w est *gagnant* s'il est possible d'atteindre le mot vide depuis w. On appelle W le langage des mots gagnants.

Par exemple, le mot $w = abbcccccccbbbaaaa$ est gagnant :

$$abbcccccccbbbaaaa \to abbbbbaaaa \to aaaaa \to \varepsilon$$

Par contre, il serait illégal, au premier tour, de supprimer le a initial (pas suffisamment long), ou seulement trois c (pas un facteur maximal).

(a) Montrer que les stratégies suivantes ne permettent pas en général d'atteindre le mot vide depuis un mot $w \in W$: [1]

  (i) supprimer le bloc de longueur au moins 3 le plus à gauche,

  (ii) supprimer un bloc de longueur maximale.

(b) Montrer que W n'est pas reconnaissable. [3]

(c) Montrer que W est algébrique. (Si vous construisez un AAP ou une grammaire, donner une explication concise pourquoi ça marche. Une preuve exhaustive n'est pas exigée.) [5]

(d) W devient-il reconnaissable si $\Sigma = \{a, b\}$? [1]

2 Problèmes indécidables

Les résultats suivants, démontrés pendant le cours, pourront être utilisés sans preuve :

- Étant donné un langage $L \subseteq \Sigma^*$ algébrique, il est indécidable si $L = \Sigma^*$.
- Étant donne une instance de $\mathrm{PCP}(f, g)$, pour deux homomorphismes f, g, on peut construire des grammaires G_f, G_g telles que $\mathcal{L}(G_f) \cap \mathcal{L}(G_g) \neq \emptyset$ ssi l'instance de PCP possède une solution.
- Le langage $L_0 := \{\, a^n b^n \mid n \geq 1 \,\}$ est algébrique, et $L_1 = \{\, a^n b^n c^n \mid n \geq 1 \,\}$ n'est pas algébrique.

(a) Soit Σ un alphabet qui contient a et b mais pas $\#$. Étant donné un langage algébrique $L \subseteq \Sigma^*$, on considère $L' := L_0\#\Sigma^* \cup \Sigma^*\#L$. Prouver que L' est reconnaissable ssi $L = \Sigma^*$. [3]

(b) Étant donné un langage algébrique L, montrer qu'il est indécidable si L est un langage reconnaissable. [2]

(c) Étant donné deux langages algébriques L, L', montrer qu'il est indécidable si $L \cap L'$ est algébrique. [3]

3 Construction de langages

(a) Montrer que le complément du langage $\{\, a^n b^n c^n \mid n \geq 0 \,\}$ est algébrique. [2]

Soit $\Sigma^* = \{a, b\}$, et soit D le langage de Dyck engendré par la grammaire $S \to aSbS \mid \varepsilon$.

(b) Donner une grammaire pour le langages de préfixes de D, à savoir [3]

$$P := \{\, u \in \Sigma^* \mid \exists v \in \Sigma^* : uv \in D \,\}.$$

(c) Donner une grammaire pour le langages de facteurs de D, à savoir [3]

$$F := \{\, v \in \Sigma^* \mid \exists u, w \in \Sigma^* : uvw \in D \,\}.$$

Note : Il faut montrer que les grammaires engendrent bien les langages demandés, mais les preuves ne devraient pas excéder quelques paragraphes.