Langages Formels

Licence Informatique/Math-Info – ENS Paris-Saclay

Examen du 23 mai 2023

Toutes les réponses devront être correctement justifiées. Temps : 2h.

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 $\langle Q, \Sigma, Z, T, q_0z_0, F \rangle$, 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_0z_0 la configuration initiale, et $F \subseteq Q$. Un AAP est *simple* s'il n'a qu'un seul état ; dans ce cas nous oublions les états dans la notation.

1 Grammaires LL

Encore pour rappel, étant donné une grammaire $G = \langle \Sigma, V, P, S \rangle$, un analyseur descendant (*top-down*) pour G est un automate à pile qui accepte $\mathcal{L}(G)$ en produisant une dérivation gauche. Un exemple (non-déterministe) d'un tel analyseur est l'automate simple et acceptant par pile vide $\mathcal{A}_G = \langle \Sigma, \Sigma \cup V, T, S \rangle$, avec deux types d'actions dans T :

- gonfler : $\langle X, \varepsilon, \alpha \rangle \in T$ pour tout $(X \rightarrow \alpha) \in P$;
- consommer : $\langle a, a, \varepsilon \rangle \in T$ pour tout $a \in \Sigma$.

Ici, le sommet de pile est à gauche.

Tout comme pour l'analyse ascendante, il existe des améliorations qui arrivent à construire des AAP déterministes pour une sous-classe des grammaires, à l'aide d'un *lookahead* de k lettres. Cela motive la définition suivante :

G est $\text{LL}(k)$ si pour toute dérivation gauche $S \rightarrow_g uX\delta$ de G avec $u \in \Sigma^*$, $X \in V$ et $\delta \in (\Sigma \cup V)^*$ et pour toute paire de productions distinctes $X \rightarrow \alpha$, $X \rightarrow \beta$, on a $\text{First}_k(\alpha\delta) \cap \text{First}_k(\beta\delta) = \emptyset$.

Intuitivement, G est $\text{LL}(k)$ si un analyseur descendant ayant X comme sommet de pile pourra choisir une production pour X en fonction du préfixe déjà consommé et les k prochaines lettres. Une grammaire est LL s'il existe k tel qu'elle est $\text{LL}(k)$.

Soit G_1 la grammaire suivante : $S \rightarrow aSb \mid ab$.

- (i) Trouver le plus petit k tel que G_1 est $\text{LL}(k)$ (s'il existe).
- (ii) Montrer que toute grammaire ambiguë n'est pas LL .
- (iii) Montrer que tout langage rationnel est engendré par une grammaire $\text{LL}(1)$.
- (iv) Trouver une grammaire non ambiguë qui n'est pas LL .

2 Mélange de mots

Soit Σ un alphabet et $u, v \in \Sigma^*$ deux mots. Le mélange de u et v est le langage

$$u \sqcup v = \{ u_1 v_1 \cdots u_n v_n \mid n \geq 0, \forall i : u_i, v_i \in \Sigma^*, u = u_1 \cdots u_n, v = v_1 \cdots v_n \}.$$

Pour deux langages $K, L \subseteq \Sigma^*$ on note $K \sqcup L = \bigcup_{u \in K, v \in L} u \sqcup v$.

Soient $a, b \in \Sigma$. Notons $K_{a,b} = \{ u \in \{a, b\}^* \mid |u|_a = |u|_b \}$ et $L_{a,b} = \{ a^n b^n \mid n \geq 0 \}$. Ces langages sont connus comme étant algébriques.

- (i) $K_{a,b} \sqcup K_{c,d}$, est-il algébrique ?
- (ii) $K_{a,b} \sqcup K_{b,c}$, est-il algébrique ?
- (iii) $L_{a,b} \sqcup L_{b,c}$, est-il algébrique ?

3 Automates à pile boustrophédon

Un automate à pile boustrophédon (AAPB) peut, dans chaque transition, choisir de déplacer la tête de lecture vers la gauche ou vers la droite ou encore la laisser sur place. De plus, le mot en entrée sera de la forme $\langle w \rangle$, où $\langle, \rangle$ sont des symboles spéciaux qui marque le début et fin du mot. On considère que w est accepté s'il existe un calcul sur $\langle w \rangle$ qui l'accepte.

De façon équivalente, un AAPB est une machine de Turing qui ne peut pas modifier sa bande d'entrée et qui a une seule bande de travail utilisée comme une pile.

- (i) Montrer qu'il existe un AAPB qui accepte $\{ a^n b^n c^n \mid n > 0 \}$.
- (ii) Montrer qu'il existe un AAPB déterministe qui accepte $\{ ww \mid w \in \{a, b\}^* \}$.

4 Quotients

Soient $K, L \subseteq \Sigma^*$ deux langages et

$$C(K, L) := \{ u \in \Sigma^* \mid \forall v \in L : uv \in K \}$$

- (i) Si K est rationnel et L algébrique, $C(K, L)$ est-il rationnel ? algébrique ?
- (ii) Si K est algébrique et L rationnel, $C(K, L)$ est-il rationnel ? algébrique ?