Grammaires algébriques multiples

Devoir à rendre au plus tard le 10 avril

Le but de ce devoir est l'étude des grammaires algébriques multiples. L'idée est de definir une classe intermédiaire de langages entre les langages algébriques et les langages contextuels, de sorte à étendre le pouvoir d'expression des grammaires algébriques tout en conservant au maximum une complexité algorithmique faible. La première partie pose la définition des grammaires algébriques multiples et en explore l'expressivité, ainsi que quelques formes normales. La deuxième partie établit les bornes de complexité des divers problèmes relatifs à cette classe de langages. Finalement, la dernière partie s'intéresse aux propriétés de fermeture de ces grammaires.

1 Préliminaires

Définition 1 (Grammaire algébrique multiple)**.** Une grammaire algébrique multiple est un tuple $G = (\Sigma, V, P, S)$ tel que :

- Σ est l'alphabet terminal.
- V est l'alphabet des symboles non-terminaux avec arité.
- S est un non-terminal d'arité 1.
- P est un ensemble de règles de production de la forme :

$$A_0(s_1, \ldots, s_{k_0}) \leftarrow A_1(x_1^1, \ldots, x_{k_1}^1) \ldots A_n(x_1^n, \ldots, x_{k_n}^n)$$

  où :
  - Les A_i sont des non-terminaux d'arité k_i.
  - Les x_j^i sont deux à deux distincts.
  - $s_\ell \in (\Sigma \cup X)^*$, avec $X = \{x_j^i \mid i \leq n, j \leq k_i\}$.
  - Chaque x_j^i apparaît au plus une fois dans le mot $s = s_1 \cdot \ldots \cdot s_{k_0}$, appelé membre gauche de la règle.
  - On appelle membre droit de la règle le mot $x_1^1 \cdot \ldots \cdot x_{k_1}^1 \cdot \ldots \cdot x_1^n \cdot \ldots \cdot x_{k_n}^n$.

Les dérivations de cette grammaire sont définies comme suit :

- Les membres gauches des règles qui n'ont pas de membres droits sont dits dérivables.
- Si $A_1(w_1^1 \ldots w_{k_1}^1), \ldots, A_n(w_1^n \ldots w_{k_n}^n)$ sont dérivables pour certains mots w_j^i et que $A_0(s_1, \ldots, s_{k_0}) \leftarrow A_1(x_1^1 \ldots x_{k_1}^1) \ldots A_n(x_1^n \ldots x_{k_n}^n)$ est une règle de production, alors $A_0(s'_1, \ldots, s'_{k_0})$ est aussi dérivable, où s'_ℓ est s_ℓ dans lequel les x_j^i sont remplacés par w_j^i.
- Le langage engendré par G est $L(G) = \{w \mid S(w) \text{ est dérivable}\}$.

1. Soit G la grammaire définie par les règles de production suivantes :

$$\begin{aligned} A(\varepsilon, \varepsilon) &\leftarrow \\ A(xa, yb) &\leftarrow A(x, y) \\ S(x \cdot y) &\leftarrow A(x, y) \end{aligned}$$

   Montrer que $L(G) = \{a^n b^n \mid n \in \mathbb{N}\}$.

2. Rappeler une démonstration du fait que $\{a^n b^n c^n \mid n \in \mathbb{N}\}$ n'est pas algébrique. Montrer que $\{a^n b^n c^n \mid n \in \mathbb{N}\}$ est un langage algébrique multiple.

   Même question pour $\{ww \mid w \in \Sigma^*\}$.

3. Montrer que tout langage algébrique est un langage algébrique multiple.

   En déduire que les grammaires algébriques multiples sont strictement plus expressives que les grammaires algébriques.

Définition 2 (Grammaire non-effaçante)**.** Une grammaire algébrique multiple est dite *non-effaçante* si, pour toutes les règles de production de la grammaire, toutes les variables x^i_j du membre droit de la règle apparaissent exactement une fois dans le membre gauche de la règle.

4. Montrer que pour toute grammaire algébrique multiple G, il existe une grammaire algébrique multiple équivalente G' telle que G' soit non-effaçante.

Définition 3 (Grammaire non-permutante)**.** Une grammaire algébrique multiple est dite *non-permutante* si, pour toutes les règles de production de la grammaire, pour tous i et tous $j_1 < j_2$, si $x^i_{j_1}$ et $x^i_{j_2}$ apparaissent dans le membre gauche de la règle, alors $x^i_{j_1}$ y apparaît avant $x^i_{j_2}$.

5. Montrer que pour toute grammaire algébrique multiple G, il existe une grammaire algébrique multiple équivalente G' telle que G' soit non-permutante. Montrer de plus que si G est non-effaçante, alors G' est aussi non-effaçante.
6. Soit $L = \{(a^n b)^k \mid n, k \in \mathbb{N}^*\}$. Montrer que L est contextuel. Montrer que L n'est pas un langage algébrique multiple.
7. Montrer que tout langage algébrique multiple est un langage contextuel.

   En déduire que les grammaires algébriques multiples sont strictement moins expressives que les grammaires contextuelles.
8. Pour cette question, on suppose que $\Sigma = \{a\}$. Montrer que, dans ce cas, les langages algébriques et les langages algébriques multiples ont le même pouvoir d'expression.

2 Problèmes de décision

1. Montrer que le problème du vide d'un langage algébrique multiple est décidable en temps linéaire.
2. Montrer que le problème de l'inclusion de deux langages algébriques multiples est indécidable.
3. Soit G une grammaire algébrique multiple fixée. Soit $w \in \Sigma^*$. Montrer que l'on peut décider si $w \in L(G)$ en temps polynomial dans la taille de w.

3 Propriétés de fermeture

1. Montrer que les langages algébriques multiples sont clos par union, concaténation et itération.
2. Montrer que les langages algébriques multiples sont clos par intersection avec un langage rationnel.
3. Montrer que les langages algébriques multiples sont clos par morphisme et morphisme inverse.
4. Montrer que les langages algébriques multiples ne sont pas clos par intersection et par complémentaire.