TELECOM Nancy — Mathématiques Appliquées pour l'Informatique

Théorie des langages : grammaires algébriques, langages algébriques

Définitions : grammaire algébrique, dérivation, arbre syntaxique d'un mot, langage engendré par une grammaire algébrique, grammaire ambiguë, propriétés

Exercice 1 Soit l'expression arithmétique suivante en notation infixée-préfixée :

$$(a + x) \ * \ f(x/b, \ s(a)) \ - \ s(f(x, y))$$

Ecrire cette expression sous forme préfixée, postfixée, fonctionnelle et arborescente.

Exercice 2 : Le but de cet exercice est d'écrire des grammaires de plus en plus correctes engendrant les expressions arithmétiques en notation infixée.

1. On se limite dans un premier temps aux deux opérateurs $+$ et $*$. Déterminer les règles de la grammaire algébrique $G_1 = (\{E\}, \{x, y, z, +, *, (,)\}, \rightarrow, E)$ engendrant les expressions arithmétiques formées à l'aide des variables x, y, z, des symboles d'opérations $+$, $*$ et des parenthèses (, et), cette grammaire n'utilisant que le seul non-terminal E.
   (a) Montrer que la grammaire obtenue engendre le mot $x \ + \ y \ * \ (x + z)$ en mettant en évidence une ou plusieurs dérivations.
   (b) Construire le ou les arbres syntaxiques des expressions suivantes : $\alpha_1 \ = \ x + (y * z)$, $\alpha_2 \ = \ x + y * z$. Que constatez-vous et quel problème met-on en évidence ?
   (c) Ecrire une grammaire G_2 engendrant les mêmes expressions mais tenant compte de l'ordre de priorité des opérateurs $+$ et $*$. Reprendre la question (b) ci-dessus et commenter les différences constatées.
2. Compléter la grammaire de la question 1.(c) de sorte que le langage engendré inclue aussi l'opérateur unaire $-$.
3. Compléter la grammaire de la question 1.(c) de sorte que le langage engendré inclue aussi l'opérateur unaire $-$ et les opérateurs binaires $-$ et $/$.
4. A l'aide de la grammaire du 3. former les arbres syntaxiques des expressions suivantes : $\alpha_3 \ = \ x - y + z$, $\alpha_4 \ = \ x \ / \ y \ * \ z$.
   Un nouveau problème est mis en évidence. Lequel ?
5. Définir une nouvelle grammaire tenant compte de l'associativité "gauche $\rightarrow$ droite" des opérateurs. Voyez-vous un nouveau problème surgir ?
6. Ajouter à la grammaire précédente un opérateur d'exponentiation $\uparrow$ ($x \uparrow y \ = \ x^y$) que l'on parenthésera "droite $\rightarrow$ gauche".

Exercice 3 Soit la grammaire $G = (\{X\}, \ \{a, b\}, \ \rightarrow, \ X)$ telle que $X \rightarrow aXbX \mid bXaX \mid \varepsilon$.

Montrer que $L(G) = \{\alpha, \ \alpha \in \{a, b\}^* \ et \ |\alpha|_a = |\alpha|_b\}$ où $|\alpha|_l$ est le nombre d'occurrences de la lettre l dans le mot α.

Exercice 4 Dans cet exercice on étudie les propriétés de stabilité des langages réguliers et des langages algébriques par rapport aux opérations sur les langages.

1. Soient L_1, L_2 et L trois langages réguliers.
   (a) Que peut-on dire de $L_1 \cup L_2$? de L_1L_2 ? de L^* ?
   (b) Montrer que $\overline{L}$ est un langage régulier. Indication : soit $\mathcal{A} = (A, Q, q_0, \delta, T)$ un automate déterministe reconnaissant le langage L, construire un automate $\overline{\mathcal{A}}$ reconnaissant $\overline{L}$.
   (c) Soient $\mathcal{A}_1 = (A, Q_1, q_0^1, \delta_1, T_1)$ et $\mathcal{A}_2 = (A, Q_2, q_0^2, \delta_2, T_2)$ deux automates finis déterministes reconnaissant respectivement L_1 et L_2, montrer que $L_1 \cap L_2$ est un langage régulier en construisant un automate fini déterministe reconnaissant $L_1 \cap L_2$.
   (d) Montrer que $\widetilde{L} = \{\tilde{\alpha}, \ \alpha \in L\}$ est régulier ($\tilde{\alpha}$ est l'inverse (ou le mot miroir) de α).
2. Soient L_1 et L_2 deux langages algébriques définis respectivement par les grammaires $G_1 = (N_1, A_1, \rightarrow_1, X_1)$ et $G_2 = (N_2, A_2, \rightarrow_2, X_2)$, et tels que $N_1 \cap N_2 = \emptyset$.
   (a) Montrer que $L_1 \cup L_2$ est un langage algébrique.
   (b) Même question pour le langage L_1L_2.

(c) Même question pour le langage L_1^*.

(d) Même question pour le langage $\widetilde{L_1}$.

Les langages algébriques ne sont pas stables par intersection, ni par passage au complémentaire.

3. **Résultats.** Langages réguliers et langages algébriques.
   On peut démontrer que l'intersection d'un langage régulier et d'un langage algébrique est un langage algébrique. Dans [1], une méthode est décrite, qui étant donnés :
   — un langage régulier et un automate qui le reconnaît
   — un langage algébrique et une grammaire qui l'engendre
   construit une grammaire algébrique qui engendre l'intersection des deux langages.

Exercice 5 Montrer que chacun des langages suivants est algébrique en donnant une grammaire qui l'engendre. On demande une justification, mais pas de démonstration formelle.

1. $L_1 = \{a^n b^n, n \in \mathbb{N}\}$
2. $L_2 = \{a^n b^m c^n d^p,\ (n, m, p) \in \mathbb{N}^3\}$
3. $L_3 = \{a^n b^m,\ (n, m) \in \mathbb{N}^2\ et\ n \neq m\}$
4. $L_4 = \{\alpha,\ \alpha \in \{a, b\}^*\ et\ |\alpha|\ pair\}$
5. $L_5 = b^*aa^*bc(a + b + c)^*(a + b)^*$
6. $L_6 = \{a^m b^n c^p,\ (m, n, p) \in \mathbb{N}^3\ et\ n > m + p\}$

Exercice 6 Nous avons rencontré cinq caractérisations des langages réguliers :

1. les langages obtenus à partir des langages finis par un nombre fini d'opérations d'**union**, de **concaténation** et de **fermeture itérative** (c'est la définition d'un langage régulier)
2. les langages dénotés par des expressions rationnelles
3. les langages reconnus par des automates finis (déterministes ou indéterministes)
4. les langages dont l'ensemble des quotients gauches est fini
5. les langages engendrés par des grammaires régulières à droites

Pour montrer qu'un langage n'est par régulier on utilise généralement les lemmes de l'étoile et un raisonnement par l'absurde.

Questions

1. Soit le langage $L_1 = \{a^n b^n,\ n \in \mathbb{N}\}$, L_1 est-il régulier ?
2. Même question pour $L_2 = \{\alpha,\ \alpha \in \{a, b\}^*\ et\ \alpha = \tilde{\alpha}\}$.
3. Même question pour $L_3 = \{\alpha\tilde{\alpha},\ \alpha \in \{a, b\}^*\}$.
4. Montrer que $L_4 = \{\alpha,\ \alpha \in \{a, b\}^*\ et\ |\alpha|_a = |\alpha|_b\}$ n'est pas régulier. Indications : répondre aux questions suivantes :
   (a) Montrer que $L = a^*b^*$ est régulier.
   (b) Quel est le langage $L_4 \cap L$? En déduire le résultat.
5. Soit $L_5 = \{a^{n^2},\ n \in \mathbb{N}\}$, L_5 est-il régulier ?
6. Même question pour $L_6 = \{a^n,\ n \in \mathbb{N}\ et\ n\ premier\}$.
7. L'affirmation suivante est-elle vraie : "tout sous-ensemble d'un langage régulier est régulier" ?

1. Pierre Marchand, Mathématiques Discrètes, Automates, langages, logique et décidabilité. Dunod