Informatique théorique III

Série 7

Prof. Nestmann, 2004

1. Grammaires quelconques

Pour chacun des langages suivants, donner une grammaire le générant. Indiquer le type des grammaires trouvées.

1. $L_0 = \{\text{airbag}, \text{mandragore}\}$, langage sur l'alphabet $\Sigma_0 = \{$a,b,...,z$\}$.
2. $L_1 = \{a^n b^n \mid n \in \mathbb{N}\}$, langage sur l'alphabet $\Sigma_1 = \{a, b\}$.
3. $L_2 = \{a^n b^n c^n \mid n \in \mathbb{N}\}$, langage sur l'alphabet $\Sigma_2 = \{a, b, c\}$.
4. $L_3 = \left\{a^{(2^n)} \mid n \in \mathbb{N}\right\}$, langage sur l'alphabet $\Sigma_3 = \{a\}$.

2. Associativité et grammaires non-contextuelles

On considère comme alphabet $\Sigma = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, -\}$.
Soit $G = (\{S, N\}, \Sigma, P, S)$ avec P tel que :

$$\begin{aligned} S &\to N - S \;\big|\; N \\ N &\to 0 \;\big|\; 1 \;\big|\; 2 \;\big|\; 3 \;\big|\; 4 \;\big|\; 5 \;\big|\; 6 \;\big|\; 7 \;\big|\; 8 \;\big|\; 9 \end{aligned}$$

On interprète les symboles $0, \ldots, 9$ par les entiers correspondants. Étant donné un arbre d'analyse (t, ψ) de G, on définit l'interprétation $I(t, \psi)$ de l'arbre, qui calcule la valeur de l'expression arithmétique que celui-ci représente, par :

$$\begin{aligned} I((n), \psi) &\triangleq \psi(n) \\ I((n\ t), \psi) &\triangleq I(t, \psi) \\ I((n\ t_1 t_2 t_3), \psi) &\triangleq I(t_1, \psi) - I(t_3, \psi) \end{aligned}$$

1. Donner l'arbre (les arbres ?) de dérivation (t, ψ) de la chaîne $1 - 2 - 3$ dans G. Que vaut $I(t, \psi)$?
2. Comme vous avez du le remarquer, l'expression précédente ne s'évalue pas en ce que l'on souhaiterait — à savoir -4. Donner alors une grammaire G' tel que l'opérateur $-$ soit *associatif à gauche* et non à droite comme G le fait.
3. Donner l'arbre de dérivation (t', ψ') de la chaîne $1 - 2 - 3$ dans G' et vérifier que $I(t', \psi') = -4$.

3. De l'ambiguïté

Soit $\Sigma = \{0, \ldots, 9, +, *\}$ un alphabet terminal et $G = (\{E, N\}, \Sigma, P, E)$ avec P tel que :

$$\begin{aligned} E &\to N \;\big|\; E + E \;\big|\; E * E \\ N &\to 0 \;\big|\; 1 \;\big|\; 2 \;\big|\; 3 \;\big|\; 4 \;\big|\; 5 \;\big|\; 6 \;\big|\; 7 \;\big|\; 8 \;\big|\; 9 \end{aligned}$$

1. Démontrer l'ambiguïté de G en exhibant deux arbres d'analyse pour le mot 1+2∗3. Le mot $45*2$ appartient-il à $\mathrm{L}(G)$?
2. Dans ce qui suit, on confond les lettres $0, \ldots, 9$ avec les chiffres qui leur correspondent et les lettres $*$ et $+$ avec les opérations idoines sur les nombres naturels.

Étant donné un arbre d'analyse (t, ψ) de G, on définit l'interprétation $I(t, \psi)$ de l'arbre, qui calcule la valeur de l'expression arithmétique que celui-ci représente, par :

$$\begin{aligned} I\big((n\ t_1\ t_2\ t_3), \psi\big) &= I(t_1, \psi)\ \mathbf{op}\ I(t_3, \psi) \qquad \text{avec } \mathbf{op} = \psi\,(\text{racine}(t_2)) \\ I\big((n\ t), \psi\big) &= I(t, \psi) \\ I\big((n), \psi\big) &= \psi(n) \end{aligned}$$

N.B. la fonction ci-dessus est bien définie sur les arbres d'analyse de G car leurs nœuds n'ont que un ou trois sous-arbres directs (pourquoi ?).

Calculer l'interprétation des deux arbres d'analyse du point précédent.

3. Changer la grammaire G en une grammaire $G' = (V, \Sigma, P', E)$ pour la rendre non-ambiguë et telle que l'ordre de priorité des opérateurs soit, du plus faible au plus fort, $+, *$ (c.-à.d, l'interprétation de $1{+}2{*}3$ devrait être 7).
4. Donner l'unique arbre d'analyse de $1{+}2{*}3$ dans G'.

   En définissant une fonction d'interprétation adéquate sur les arbres d'analyse de G' comme ci-dessus, l'unique interprétation de $1{+}2{*}3$ devrait être 7. Cependant, parfois, on aimerait pouvoir écrire des mots tels que $(1{+}2){*}3$, à interpréter par 9, les parenthèses servant à changer l'ordre de priorité des opérateurs.

   Donner une grammaire $G'' = (V', \Sigma \cup \{), (\}, P'', E)$ qui garde l'ordre de priorité des opérateurs de la grammaire G' mais qui permet aussi d'introduire des parenthèses pour changer l'ordre de priorité des opérateurs.
5. Soit la grammaire $G''' = (\{E, N\}, \Sigma, P''', E)$ avec P''' tel que :

$$\begin{aligned} E &\to N \ \big|\ +EE\ \big|\ *EE \\ N &\to 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9 \end{aligned}$$

   Cette grammaire génère le même genre d'expression que G mais en notation préfixe. Donner les arbres de dérivation de $+1{*}23$ et $*{+}123$ et démontrer que cette grammaire n'est pas ambiguë.

   Pour cela utiliser le lemme suivant :

   Lemme 3.1 *Soit G une grammaire non-contextuelle et $u\beta \Rightarrow^* w$ une dérivation dans G avec $u, w \in \Sigma^*$. Alors il existe $w' \in \Sigma^*$ tel que $w = u \cdot w'$.*

4. De l'ambiguïté (bis)

Donner une grammaire G' ambiguë qui génère le même langage que la grammaire $G = (\{S, S'\}, \{), (\}, P, S)$ avec P tel que :

$$\begin{aligned} S &\to S'S \ \big|\ \epsilon \\ S' &\to (S) \end{aligned}$$

Démontrer l'ambiguïté de G'.