Grammaires et analyse - Semaine du 18/11/2002
Formes Normales

Yan Jurski, Aurelian Lavric, Dominique Poulalhon

1 Forme normale de Chomsky

On dit qu'une grammaire est sous *forme normale de Chomsky* (ou *forme normale quadratique* si toutes les productions qu'elle contient sont sous la forme

$$V \to WX$$

ou

$$V \to a$$

où V, W, X sont des symboles non-terminaux, et a est un terminal.

Dans les grammaires suivantes, effectuez la suite de transformations suivante : élimination des symboles inutiles, élimination des ε-productions, élimination des productions unaires, mise sous forme normale de Chomsky.

$$S \to aS \mid Sb \mid c \qquad \begin{array}{l} S \to A \\ A \to b \mid c \end{array} \qquad S \to abcdef \quad \begin{array}{l} S \to aS \mid Ab \\ A \to cB \\ B \to dA \mid e \end{array}$$

$$\begin{array}{l} S \to aSAb \mid bSS \mid d \\ A \to cASaA \mid bcd \end{array} \quad S \to x \mid lSr \mid SpS \mid SmS$$

2 Forme normale de Greibach

Vérifiez que les grammaires suivantes sont sous forme normale de Chomsky et mettez-les sous forme normale de Greibach.

$$\begin{array}{l} A_1 \to A_2A_3 \\ A_2 \to a \mid A_1A_2 \\ A_3 \to b \end{array} \quad \begin{array}{l} A_1 \to A_2A_3 \\ A_2 \to a \mid A_1A_3 \\ A_3 \to b \mid A_2A_3 \end{array} \quad \begin{array}{l} S \to a \mid AA \\ A \to b \mid SS \end{array}$$

3 Grammaires linéaires

Une grammaire est dite *linéaire* si dans tout membre droit de production il y a au plus un symbole non terminal. Un langage est linéaire lorsqu'il existe une grammaire linéaire qui l'engendre.

Montrez que le langage $\{a^nb^n \mid n \in \mathbf{N}\}$ est linéaire. Qu'en est-il du langage $\{u\tilde{u} \mid u \in \{a,b\}^*\}$? Et des langages reguliers ?

Vérifiez que toutes les transformations vues jusqu'ici (élimination des symboles inutiles, élimination des ε-productions, élimination des productions unaires) préservent la linéarité.

Remarque : la classe des langages linéaires est close par intersection et union avec un langage rationnel. On peut montrer que le langage $\{a^nb^nc^pd^p \mid n,p\}$ n'est pas linéaire.

4 Ambiguité

On rappelle qu'une grammaire est ambigûe lorsqu'il existe deux arbres de dérivation distincts pour un mot donné.

1. Montrez que les grammaires suivantes sont ambigües :

$$S \to \varepsilon \mid aSb \mid aaSb \qquad S \to x \mid SpS \mid SmS$$

On dit qu'un langage est *intrinsuement ambigu* lorsque toute grammaire qui le génère est ambigüe.

2. Montrez que l'union de deux langages non-intrinsèquement ambigus disjoints n'est pas intrinsèquement ambigüe.

3. Soit G une grammaire dont toutes les productions ont la forme

$$U \to a\alpha$$

où a est un terminal et α une suite de symboles (terminaux ou non) telle que dès qu'on a

$$U \to a\alpha$$

et

$$U \to a\alpha'$$

alors $\alpha = \alpha'$. Montrez que G n'est pas ambigüe.

3. Montrez qu'un langage rationnel n'est pas intrinsèquement ambigu.

5 Automates à pile

Construisez un automate à pile à un seul état qui reconnaît à pile vide le langage engendré par la grammaire suivante :

$$S \to +SS \mid \times SS \mid x$$

Cet automate est-il déterministe ?

Même question pour les grammaires de l'exercice précédent.