

Examen

Module: Théorie des Langages (INF202)

Pierre SENELLART

pierre.senellart@telecom-paristech.fr

29 septembre 2008

1 Grammaire simplifiée pour XML (4 points)

Soit $\Sigma = \{a, <, >, /\}$. On considère la grammaire hors-contexte G suivante sur l'alphabet Σ , qui décrit une version simplifiée du langage de balisage XML :

$$S \rightarrow \langle a \rangle T \langle /a \rangle \mid \langle a / \rangle$$

$$T \rightarrow aT \mid ST \mid \varepsilon$$

1. (2 points) Donner un arbre de dérivation pour chacun des deux mots suivants :

$$\langle a \rangle \langle a / \rangle aa \langle /a \rangle$$

$$\langle a \rangle aa \langle a \rangle \langle /a \rangle \langle /a \rangle$$

2. (2 points) Le mot $\langle a \rangle \langle a / \rangle \langle /a \rangle \langle /a \rangle$ est-il engendré par la grammaire ? Et le mot $a \langle a / \rangle a$? Justifier vos affirmations.

2 Grammaires linéaires (12 points)

1. On considère la grammaire G suivante, sur l'alphabet $\Sigma = \{a, b, c, d\}$:

$$S \rightarrow abA$$

$$A \rightarrow cA \mid adB$$

$$B \rightarrow A \mid b$$

- (1 point) Quel est le type de G dans la hiérarchie de Chomsky ?
- (2 points) Montrer que G engendre un langage rationnel en construisant une grammaire *régulière* G' équivalente à G .
- (2 points) Construire un automate reconnaissant le langage engendré par G et G' en utilisant une technique vue en cours.
- (2 points) En déduire une expression rationnelle décrivant le langage engendré par G et G' en utilisant une technique vue en cours.

- e) (1 point) Simplifier cette expression rationnelle en utilisant les abréviations $u?$, u^+ , etc.
2. (3 points) On appelle *grammaire linéaire à droite* une grammaire dans laquelle toute règle de production $\alpha \rightarrow \beta$ est soit de la forme $A \rightarrow uB$ avec u dans Σ^* et A, B dans N , soit de la forme $A \rightarrow u$ avec u dans Σ^* . S'inspirer de la question précédente pour montrer que toute grammaire linéaire à droite est équivalente à une grammaire régulière.
3. (1 points) On appelle *grammaire linéaire* une grammaire dans laquelle toute règle de production $\alpha \rightarrow \beta$ est soit de la forme $A \rightarrow uBv$ avec u, v dans Σ^* et A, B dans N , soit de la forme $A \rightarrow u$ avec u dans Σ^* . Les langages engendrés par une grammaire linéaire sont-ils toujours rationnels?

3 Langage des nombres premiers (4 points)

Soit $\Sigma = \{a\}$. On considère le langage $L = \{a^2, a^3, a^5, a^7, a^{11} \dots\}$, constitué des mots de Σ^* ayant un nombre *premier* de a .

1. (3 points) Montrer que L n'est pas hors-contexte.
2. (1 point) Sans essayer de prouver quoi que ce soit, discuter de la position possible de L dans la hiérarchie de Chomsky.