

A.1 Grammaires formelles

1. Soit la grammaire $\mathcal{G}_2 = \langle \{ (,) \}, \{ S \}, S, \{ S \rightarrow \varepsilon \mid (S)S \} \rangle$. Quel est le langage engendré par cette grammaire ? Soit le mot $((()((())))$. Donner la dérivation gauche, la dérivation droite, et l'arbre syntaxique pour ce mot.
2. Soit la grammaire $S \rightarrow aSb \mid \varepsilon$. Quel est le langage reconnu par cette grammaire ? Comment rendre cette grammaire ε -libre ?
3. Donner une grammaire qui reconnaisse le langage L (alphabet $X = \{a, b, c\}$).

$$L = \{w \in X^* / w = a^n b^n c^p; n > 0 \text{ et } p > 0\}$$

4. Soient les deux grammaires suivantes. Pour chacune d'entre elles, donnez le langage engendré, et indiquez le type de la grammaire dans la classification de Chomsky. Commentez brièvement.

$S \rightarrow S_1 S_2$	$S \rightarrow aSBC$
$S_1 \rightarrow aS_1 b \mid ab$	$S \rightarrow aBC$
$S_2 \rightarrow cS_2 \mid c$	$CB \rightarrow BC$
	$aB \rightarrow ab$
	$bB \rightarrow bb$
	$bC \rightarrow bc$
	$cC \rightarrow cc$

5. Un groupe nominal est soit un déterminant suivi d'un nom soit un groupe nominal suivi d'un groupe prépositionnel. Un groupe prépositionnel est une préposition suivie d'un groupe nominal.
 - (a) Ecrire la grammaire G du groupe nominal
 - (b) Ecrire les deux arbres de dérivation que G associe au mot D N P D N P D N où D est le symbole associé à la catégorie "déterminant", N le symbole associé à la catégorie "nom" et P celui associé à la catégorie "préposition".
 - (c) Combien d'arbres de dérivation admet le mot D N P D N P D N P D N ?
 - (d) Parmi les deux arbres de la question (b), lequel choisiriez vous pour représenter la structure du groupe nominal "le chat de la voisine de la concierge" ?
 - (e) Modifiez la grammaire G de sorte qu'elle n'associe au mot D N P D N P D N que l'analyse que vous avez identifiée dans la question précédente.

6. Les deux grammaires suivantes reconnaissent un sous-ensemble des expressions arithmétiques habituelles (par exemple, le mot $a + b \times c$). Quelle est la différence entre les deux ?

$E \rightarrow E + E \mid E \times E$	$E \rightarrow E + T \mid T$
$E \rightarrow (E) \mid a \mid b \mid c$	$T \rightarrow T \times F \mid F$
	$F \rightarrow (E) \mid a \mid b \mid c$

7. Soit la grammaire régulière suivante :

$$G = \langle \{a, b, c, d\}, \{S, T, V\}, S, \left\{ \begin{array}{l} S \rightarrow aS|bS|cT, \\ T \rightarrow aV|bV|aT|bT|cT|dT, \\ V \rightarrow \varepsilon \end{array} \right\} \rangle.$$

- a. Dessiner un automate (non déterministe) reconnaissant $L_G(S)$.
 - b. Déterminer cet automate.
8. Soit la grammaire suivante : $\mathcal{G}_1 = \langle \{a, b\}, \{S\}, S, \{ S \rightarrow aSbS|bSaS|\varepsilon \} \rangle$.
 - (a) Proposer une grammaire \mathcal{G}_2 ε -libre qui reconnaît le même langage.

- (b) Dessiner les deux arbres de dérivation qui correspondent à l'analyse du mot *aabbaabbbaab* au moyen des deux grammaires \mathcal{G}_1 et \mathcal{G}_2 .
9. Règles simples (ou productions singulières). Transformer la grammaire suivante en grammaire sans règles simples.
 $S \rightarrow AB \mid A$; $A \rightarrow aB \mid bA \mid aSb$; $B \rightarrow S \mid b$
10. Soit la grammaire hors contexte suivante :
- | | | |
|-----|---|--|
| S | → | p |
| p | → | gn v1 <i>que</i> p gn v2 |
| gn | → | np det nc |
| np | → | <i>Léa</i> <i>Luc</i> <i>Ève</i> <i>Max</i> |
| nc | → | <i>femme</i> <i>homme</i> <i>étudiante</i> <i>étudiant</i> <i>filles</i> <i>garçon</i> |
| det | → | <i>le</i> <i>la</i> <i>l'</i> |
| v1 | → | <i>pense</i> <i>croit</i> <i>voit</i> <i>sait</i> <i>dit</i> <i>raconte</i> |
| v2 | → | <i>se promène</i> <i>marche</i> <i>part</i> |
- (a) Donner quatre phrases distinctes reconnues par cette grammaire, contenant respectivement 0, 1, 2 et 3 fois le mot *que*.
- (b) Pour quelles raisons ces phrases ne sont-elles pas toutes correctes en français ? Comment modifier la grammaire pour corriger cela ?
- (c) Donner l'arbre de dérivation de *Luc sait que la femme croit que Léa part*.
11. Donner une grammaire régulière (vocabulaire $V = \{a, b, c, d, \dots, y, z\}$) pour le langage qui contient l'ensemble des mots, de deux lettres minimum, composés en alternance d'une consonne et d'une voyelle et débutant ou finissant par une consonne et une voyelle. Donner une grammaire algébrique engendrant le même langage.
12. Proposer une grammaire algébrique qui engendre le langage $\{a^n db^p da^n, n > 0, p \geq 0\}$.
13. Soit l'alphabet $X = \{+, =, a\}$. (1) Donner une grammaire algébrique pour le langage L dont chaque mot représente une addition correcte de deux suites de caractères a . Par exemple L contient le mot $aa + aaaa = aaaaaa$. (2) Donner un automate à pile qui reconnaît le même langage.
14. Soit le langage L_1 sur le vocabulaire $V = \{l', homme, ours, qui, a, vu\}$ formé de l'ensemble des phrases finies de la forme *l'homme qui a vu l'ours*, *l'homme qui a vu l'homme qui a vu l'ours*, *l'homme qui a vu l'ours*, *l'homme qui a vu l'homme qui a vu ... qui a vu l'ours*.
- (a) Donner une grammaire algébrique (*context-free*) engendrant L_1 .
- (b) Donner une grammaire régulière engendrant L_1 .
- Soit L_2 le langage engendré par la grammaire \mathcal{G}_2 :
- | | | |
|-----|---|-----------------|
| S | → | NP Rel |
| NP | → | l'homme |
| | | l'ours |
| Rel | → | qui a vu NP Rel |
| | | qui a vu l'ours |
- (c) Mettre la grammaire \mathcal{G}_2 sous forme normale de Chomsky.
- (d) Proposer une grammaire \mathcal{G}_3 qui engendre le langage L_3 , sur-ensemble de L_2 mais dans lequel les symboles *ours* et *homme* sont **strictement** interchangeables.
- (e) Décrire informellement les différences entre les langages L_1 et L_2 .
- (f) Donner une grammaire algébrique du langage $L_2 \setminus L_1$.