

**Partiel - durée : 2h**  
**(documents issus du cours autorisés)**

---

**Exercice 1.***Hamming*

On travaille sur l'alphabet  $\{0, 1\}$ . La distance de Hamming  $H(x, y)$  entre deux mots  $x$  et  $y$  est le nombre de positions qui les distinguent, par exemple  $H(110, 011) = 2$ . Si  $|x| \neq |y|$ , alors leurs distance de Hamming est infinie. Si  $x$  est un mot et  $A$  est un langage sur  $\{0, 1\}$ , la distance de Hamming entre  $x$  et  $A$  est la distance entre  $x$  et le mot le plus proche de  $x$  dans  $A$  :

$$H(x, A) := \min_{y \in A} H(x, y)$$

Pour un langage  $A \subseteq \{0, 1\}^*$  et  $k \geq 0$ , on définit l'ensemble de mots de distance de Hamming au plus  $k$  de  $A$  :

$$N_k(A) := \{x \mid H(x, A) \leq k\}$$

Prouvez que si  $A$  est un langage régulier alors  $N_k(A)$  est régulier.

**Exercice 2.***Étoile*

Parmi les langages suivants il y a *au moins un* langage algébrique et *au moins un* non-algébrique. Choisissez deux langages tels que un est algébrique et l'autre non et démontrez le (pour montrer qu'un langage est non-algébrique vous utiliserez le lemme de l'étoile pour les langages hors-contexte) :

- $\{w \in \{a, b\}^* \text{ tels que } |w|_b = 2|w|_a + 3\}$
- $\{w\#x \mid w, x \in \{a, b\}^* \text{ et } w \text{ est un sous-mot de } x\}$ .
- $\{a^p \mid p \text{ est premier}\}$ .
- $\{a^{n_0} b a^{n_1} b \dots a^{n_k} b \mid k \geq 0 \text{ et } \exists j \geq 0, n_j \neq j\}$

**TOURNEZ SVP.**

### Exercice 3.

La grammaire de Sheila

**Définition 1.** Une grammaire hors-contexte  $G = (V, \Sigma, P, S)$  est sous forme normale de Greibach si toutes ses règles sont de la forme

i)  $A \rightarrow aB_1 \dots B_n$

ii)  $A \rightarrow a$

iii)  $S \rightarrow \epsilon$

Où  $B_1, \dots, B_n \in V - \{S\}$  et  $a \in \Sigma$ .

On l'appelle également forme m-standard lorsque  $n \leq m$  pour toutes les règles.

Nous allons prouver le théorème suivant

**Théorème 1.** Tout langage hors-contexte peut être généré par une grammaire hors-contexte sous forme normale de Greibach.

#### 1. Montrer le lemme suivant

**Lemme 1** (substitution). Soit  $G = (V, \Sigma, P, S)$  une grammaire hors-contexte,  $\pi = A \rightarrow \alpha_1 B \alpha_2$  une production de  $P$  avec  $B \rightarrow \beta_1 | \dots | \beta_r$  la règle de  $P$  avec  $B$  à gauche. Soit  $G_1 = (V, \Sigma, P_1, S)$  avec

$$P_1 = (P - \{\pi\}) \cup \{A \rightarrow \alpha_1 \beta_1 \alpha_2 | \alpha_1 \beta_2 \alpha_2 | \dots | \alpha_1 \beta_r \alpha_2\}$$

alors  $L(G_1) = L(G)$ .

#### 2. Montrer le lemme suivant

**Lemme 2** (inversion). Soit  $G = (V, \Sigma, P, S)$  une grammaire hors-contexte ne produisant pas le mot vide et

$$A \rightarrow A \alpha_1 | \dots | A \alpha_r$$

avec  $\alpha_i \neq \epsilon$  pour tout  $i$ , l'ensemble des règles avec  $A$  en tête telles que le symbole le plus à gauche du corps soit  $A$ . Soit

$$A \rightarrow \beta_1 | \dots | \beta_s$$

l'ensemble des règles restantes avec  $A$  en tête. Soit  $G_1 = (V \cup \{Z\}, \Sigma, P_1, S)$ , où  $Z$  est un nouveau symbole non-terminal et  $P_1$  l'ensemble des productions de  $P$  avec toutes les productions avec  $A$  en tête remplacées par

$$A \rightarrow \beta_1 Z | \dots | \beta_s Z | \beta_1 | \dots | \beta_s$$

$$Z \rightarrow \alpha_1 Z | \dots | \alpha_r Z | \alpha_1 | \dots | \alpha_r$$

alors  $L(G_1) = L(G)$ .

Sans perte de généralité, on partira d'une grammaire  $G = (V, \Sigma, P, S)$  sous forme normale de Chomsky, telle que  $S$  n'apparaît dans aucun membre droit. **Supposons pour le moment que  $G$  ne produit pas le mot vide.**

#### 3. Soit $V = \{A_1, \dots, A_m\}$ avec $S = A_1$ . L'ensemble des règles de production peut être classé en deux catégories

$$A_i \rightarrow BC \text{ avec } B, C \in V \setminus \{S\}$$

$$A_i \rightarrow a$$

Construisez une grammaire équivalente dont les règles de production peuvent être classées en trois catégories

- (a)  $A_i \rightarrow A_j \alpha$  avec  $i < j, \alpha \in ((V - \{S\}) \cup \{Z_1, \dots, Z_n\})^+$
- (b)  $A_i \rightarrow a \alpha$  avec  $a \in \Sigma, \alpha \in ((V - \{S\}) \cup \{Z_1, \dots, Z_n\})^*$
- (c)  $Z_i \rightarrow \alpha$  avec  $\alpha \in ((V \cup \Sigma - \{S\}) \cup \{Z_1, \dots, Z_n\})^+$

*Indication* : On pourra construire la grammaire par récurrence sur  $k$  avec la propriété que toutes les règles dont le membre gauche est  $A_i$  sont soit de la forme (a), soit de la forme (b), pour  $i \leq k$ .

4. Toujours sous l'hypothèse que le mot vide n'appartient pas au langage, conclure.
5. Conclure si de plus  $\epsilon$  est dans le langage.
6. En quoi la forme normale de Greibach est-elle intéressante ?
7. Mettre la grammaire suivante sous forme normale de Greibach

$$A_1 \rightarrow A_2 A_3$$

$$A_2 \rightarrow A_1 A_2 | 1$$

$$A_3 \rightarrow A_1 A_3 | 0$$

**TOURNEZ SVP.**

**Exercice 4.**

Analyse Syntaxique

Soit  $G = (V, \Sigma, R, S)$  une grammaire hors-contexte sans production vide. On appelle couple pointé pour un mot  $w = a_1 a_2 \dots a_n$  l'élément  $[A \rightarrow \alpha \bullet \beta, i]$  où  $i \in [0, n]$ , le point  $\bullet \notin \Sigma \cup V$  est un symbole spécial et  $A \rightarrow \alpha \beta$  est une règle de production de  $R$ .

Soit l'algorithme d'analyse syntaxique suivant :

**Algorithm 1:** Algorithme *Earley*

**Données:** Grammaire hors-contexte  $G = (V, \Sigma, R, S)$  sans production vide, un mot  $w = a_1 a_2 \dots a_n \in \Sigma^*$

**Sorties:** Des listes d'analyse  $L_0, L_1, \dots, L_n$  pour  $w$

**Résultat:**  $w \in L(G)$  si et seulement si  $[S \rightarrow \alpha \bullet, 0] \in L_n$

**début**

Construction de la liste  $L_0$  :

1 **si**  $S \rightarrow \alpha \in R$  **alors**  
     | mettre  $[S \rightarrow \bullet \alpha, 0]$  dans  $L_0$

Exécuter ensuite les étapes 2 et 3 tant qu'elles ajoutent un élément dans  $L_0$  :

2 **si**  $[B \rightarrow \gamma \bullet, 0] \in L_0$  **alors**  
     | **pour chaque** couple pointé  $[A \rightarrow \alpha \bullet B \beta, 0] \in L_0$  **faire**  
         | | mettre  $[A \rightarrow \alpha B \bullet \beta, 0]$  dans  $L_0$

3 **si**  $[A \rightarrow \alpha \bullet B \beta, 0] \in L_0$  **alors**  
     | **pour chaque** règle de  $R$  de la forme  $B \rightarrow \gamma$  **faire**  
         | | mettre  $[B \rightarrow \bullet \gamma, 0]$  dans  $L_0$

Construction des listes  $L_j$  à partir de  $L_0, L_1, \dots, L_{j-1}$  :

4 **pour chaque**  $[B \rightarrow \alpha \bullet a_j \beta, i] \in L_{j-1}$  **faire**  
     | mettre  $[B \rightarrow \alpha a_j \bullet \beta, i]$  dans  $L_j$

Exécuter ensuite les étapes 5 et 6 tant qu'elles ajoutent un élément dans  $L_j$  :

5 **si**  $[B \rightarrow \gamma \bullet, i] \in L_j$  **alors**  
     | **pour chaque** couple pointé  $[A \rightarrow \alpha \bullet B \beta, k] \in L_i$  **faire**  
         | | mettre  $[A \rightarrow \alpha B \bullet \beta, k]$  dans  $L_j$

6 **si**  $[A \rightarrow \alpha \bullet B \beta, i] \in L_j$  **alors**  
     | **pour chaque** règle de  $R$  de la forme  $B \rightarrow \gamma$  **faire**  
         | | mettre  $[B \rightarrow \bullet \gamma, j]$  dans  $L_j$

**fin**

1. Faites dérouler l'algorithme pour le mot  $(a + b) = b$  et la grammaire suivante :

$S \rightarrow (S) \mid R$

$R \rightarrow E = E$

$E \rightarrow (E + E) \mid a \mid b$

2. Expliquez qu'est-ce que signifie un élément  $[A \rightarrow \alpha \bullet \beta, i] \in L_j$ .