Exercice 1 : Soit $\Sigma = \{a, b\}$, $\mathcal{V} = \{S, D, T, X\}$, et $\mathcal{G}$ la grammaire non contextuelle décrite par les règles suivantes :

$$\begin{aligned} S &\to XSX \mid D \\ D &\to aTb \mid bTa \\ T &\to XTX \mid X \mid \varepsilon \\ X &\to a \mid b \end{aligned}$$

1. Donner 3 mots qui sont acceptés par $\mathcal{G}$.
2. Donner 3 mots qui ne sont pas acceptés par $\mathcal{G}$.
3. Est-ce que $T \Rightarrow aba$?
4. Est-ce que $T \Rightarrow^* aba$?
5. Est-ce que $T \Rightarrow T$?
6. Est-ce que $T \Rightarrow^* T$?
7. Est-ce que $XXX \Rightarrow^* aba$?
8. Est-ce que $X \Rightarrow^* aba$?
9. Est-ce que $T \Rightarrow^* XX$?
10. Est-ce que $T \Rightarrow^* XXX$?
11. Est-ce que $D \Rightarrow^* \varepsilon$?
12. Donner une description en français de $\mathcal{L}(\mathcal{G})$.

Exercice 2 : Soit $\Sigma = \{a, b, c, \ldots, z, +, (,)\}$, $\mathcal{V} = \{S, A, P, F, T\}$, et $\mathcal{G}$:

$$\begin{aligned} S &\to A \\ A &\to A + P \mid P \\ P &\to PF \mid F \\ F &\to (A) \mid T \\ T &\to a \mid b \mid c \mid \cdots \mid z \end{aligned}$$

1. Donner un arbre de dérivation de $b + (a + b)c + a(bc)$.
2. Chercher un arbre de dérivation de $a + ()$. Qu'en conclure ?

Exercice 3 : Soit $\Sigma = \{a, b\}$.

Pour chacun des langages suivants, donner une grammaire non contextuelle la reconnaissant :

1. $L_1 = \{w \in \Sigma^\star \mid |w|_a \geq 3\}$;
2. $L_2 = \{w \in \Sigma^\star \mid w \text{ commence et fini par la même lettre}\}$;
3. $L_3 = \{w \in \Sigma^\star \mid |w| \equiv 1 \mod 2\}$;
4. $L_4 = \{w \in \Sigma^\star \mid |w| \equiv 1 \mod 2 \text{ et son symbole du milieu est } a\}$;
5. $L_5 = \{w \in \Sigma^\star \mid w \text{ est un palindrome}\}$;
6. $L_6 = \overline{L_5}$;
7. $L_7 = \emptyset$;
8. $L_8 = \{w \in \Sigma^\star \mid |w|_a \geq |w|_b\}$;
9. $L_9 = \{a^n b^n \mid n \geq 0\}$;
10. $L_{10} = \overline{L_9}$;
11. $L_{11} = \{w\#x \mid w, x \in \Sigma^\star,\ w^R \text{ est un sous-mot de } x\}$;
12. $L_{12} = \{x_1\#x_2\# \cdots \#x_k \mid k \geq 1,\ \text{chaque } x_i \in \Sigma^\star,\ \exists i, j \text{ tels que } x_i = x_j^R\}$;

Exercice 4 : Soit $\Sigma = \{$ `[`,`]`, `1`, `;` $\}$.
Donner une grammaire non contextuelle reconnaissant les listes OCaml de `1`.
Exemples : `[]`, `[1]`, `[1;1]` sont des mots reconnus par la grammaire.

Exercice 5 : Soit $\Sigma = \{a, b\}$ et $\mathcal{V} = \{S\}$. On considère la grammaire $\mathcal{G}$ suivante :

$$S \to aS \mid Sb \mid a \mid b$$

1. Soit $v \in \mathcal{L}(\mathcal{G})$. Montrer par récurrence sur $|v|$ que ba n'est pas un sous-mot de v.
2. Donner $\mathcal{L}(\mathcal{G})$.

Exercice 6 : Soit $\Sigma = \{a, b, c\}$.
Donner une grammaire non contextuelle reconnaissant les langages suivants :
1. $L_1 = \{a^n b^n \mid n \in \mathbb{N}\}$.
2. $L_2 = \{a^n b^m \mid n \leq m\}$.
3. $L_3 = \{a^n b^m \mid n \geq m\}$.
4. $L_4 = \{a^i b^j c^k \mid i = j \lor j = k\}$.
5. $L_5 = \{a^i b^j c^k \mid i \neq j \lor j \neq k\}$.
6. $L_6 = \{v \in \{a, b\}^\star \mid |v|_a = |v|_b\}$.
7. $L_7 = \{v \in \{a, b\}^\star \mid |v|_a = 2|v|_b\}$.
8. $L_8 = \{v \in \{a, b\}^\star \mid v$ n'est pas de la forme ww avec $w \in \Sigma^\star\}$.

Exercice 7 (Grammaire ambigüe) : Soit $\Sigma = \{a, b\}$, $\mathcal{V} = \{S\}$, et $\mathcal{G}$:

$$S \to aS \mid aSbS \mid \varepsilon$$

1. Montrer que $\mathcal{G}$ est ambigüe.
   Indication : on pourra s'intéresser à au mot aab.
2. Montrer que $\mathcal{L}(\mathcal{G}) = \{v \in \Sigma^\star \mid \text{pour tout } u \text{ préfixe de } v,\ |u|_a \geq |u|_b\}$.
3. Donner une grammaire non ambigüe reconnaissant $\mathcal{L}(\mathcal{G})$.

Exercice 8 : Soit $\Sigma = \{0, 1\}$ et $\mathcal{V} = \{S, A, B\}$. On considère la grammaire suivante :

$$\begin{aligned} S &\to A1B \\ A &\to 0A \mid \varepsilon \\ B &\to 0B \mid 1B \mid \varepsilon \end{aligned}$$

1. Donner une dérivation à gauche et une dérivation à droite des mots suivants :
   (a) 00101.
   (b) 1001.
   (c) 00011.
2. Donner $\mathcal{L}(\mathcal{G})$.
3. Montrer que $\mathcal{G}$ n'est pas ambigüe.
4. Donner une grammaire reconnaissant $\mathcal{L}(\mathcal{G})$ mais qui est ambigüe.

Exercice 9 (Écriture préfixe) : Soit $\Sigma = \{+, *, -, x, y\}$, $\mathcal{V} = \{S\}$, et $\mathcal{G}$:

$$S \to +SS \mid *SS \mid -SS \mid x \mid y$$

1. Donner un arbre de dérivation du mot $+ * -xyxy$.
2. Montrer que $\mathcal{G}$ n'est pas ambigüe.

Exercice 10 : Soit $\Sigma = \{a, b\}$, $\mathcal{V} = \{S, T\}$, et $\mathcal{G}$:

$$S \to SS \mid T$$
$$T \to aTb \mid ab$$

1. Décrire $\mathcal{L}(\mathcal{G})$.
2. Montrer que $\mathcal{G}$ est ambigüe.
3. Donner une grammaire non ambigüe $\mathcal{H}$ telle que $\mathcal{L}(\mathcal{H}) = \mathcal{L}(\mathcal{G})$.

Exercice 11 (Stabilité) : Soit Σ un alphabet.
Montrer que l'ensemble des langages non contextuels est stable par union, concaténation, et étoile de Kleene.

Exercice 12 (Langage de Dyck) : Soit $\Sigma = \{(,)\}$. Soit D l'ensemble des mots bien parenthésés sur Σ.

1. Montrer que D n'est pas un langage régulier.
2. Donner une grammaire non contextuelle reconnaissant D.
3. Montrer que D est l'ensemble des mots v tels que :
   - $|v|_(= |v|_)$;
   - pour tout préfixe u de v, $|u|_(\geq |u|_)$.
4. Donner un programme OCaml qui vérifie qu'une chaîne de caractère sur $\Sigma = \{(,), \{, \}, [,]\}$ est un mot bien parenthésé.

Exercice 13 : Soit $\Sigma = \{a, b, (,), +, ^\star, \emptyset, e\}$.
On peut voir Σ comme l'ensemble des symboles utilisés pour écrire une expression régulière sur $\{a, b\}$, mais où e est utilisé à la place de ε (pour éviter de confondre le symbole ε et le mot vide ε), et où $+$ est utilisé à la place de $\mid$ (pour ne pas confondre le symbole $\mid$ de Σ et le symbole $\mid$ permettant d'écrire des règles d'une grammaire).

1. Donner une grammaire non contextuelle reconnaissant l'ensemble des expressions régulières syntaxiquement correctes sur $\{a, b\}$.
2. Votre grammaire est-elle ambigüe ? Si oui, en donner une qui est non ambigüe.

Exercice 14 : Soit $\Sigma = \{0, \#\}$, $\mathcal{V} = \{S, T, U\}$, et $\mathcal{G}$:

$$S \to TT \mid U$$
$$T \to 0T \mid T0 \mid \#$$
$$U \to 0U00 \mid \#$$

1. Décrire $\mathcal{L}(\mathcal{G})$ en français.
2. Montrer que $\mathcal{L}(\mathcal{G})$ n'est pas régulier.