# Exercices du cours ”outils formels de base”

## Master LTD - ENS - Paris 8

1. **Expressions régulières**
   Soit $A^*$ le monoïde libre sur l'alphabet {a, b, c}. Si $E$ et $F$ sont des sous-ensembles de $A^*$, on note $E + F$ l'union des deux ensembles et $E.F$ l'ensemble des mots de $A^*$ qui sont la concaténation d'un mot de $E$ et d'un mot de $F$. On note également $E^*$ l'ensemble des mots de $A^*$ qui sont obtenus en concaténant 0, 1, 2, ... n, ou un nombre quelconque de mots de $E$. (Si un mot est la concaténation de 0 mot de $E$, on pose que c'est le mot vide $\epsilon$). Quels langages désignent les expressions suivantes :
   - $\{b\}\{a\}^*$
   - $\{a\}^*\{b\}\{a\}^*\{b\}\{a\}^*$
   - $(\{a\} + \{b\})^*$
   - $(\{a\} + \{b\})^*(\{aa\} + \{bb\})(\{a\} + \{b\})^*$
2. Vérifier que :
   - $E.(F + G) = E.F + E.G$
   - $E^* = \{\epsilon\} + E + E^2 + E^3 + ....$
   - $E^* = \{\epsilon\} + E.E^*$
3. Un ensemble de mots peut être représenté plus simplement comme une somme. Par exemple, {a, b, c, d} sera représenté par $a+b+c+d$. Ainsi, $\{b\}\{a\}^*$ sera représenté simplement par $ba^*$, $\{a\}^*\{b\}\{a\}^*\{b\}\{a\}^*$ par $a^*ba^*ba^*$, $(\{a\} + \{b\})^*$ par $(a + b)^*$ et $(\{a\}+\{b\})^*(\{aa\}+\{bb\})(\{a\}+\{b\})^*$ par $(a+b)^*(aa+bb)(a+b)^*$. Démontrer que :
   - $(a + b)(c + d) = ac + ad + bc + bd$
   - $a(ba)^* = (ab)^*a$
4. Quels ensembles de phrases représentent les expressions régulières suivantes :
   $(the + this + a)\ girl\ (came + arrived)(early + late)$
   $a\ very^*\ beautiful\ (girl + boy)$
5. **Systèmes thuiens et semi-thuiens**
   Soit le système $R$ sur l'alphabet {a, b} tel que $ab \leftrightarrow \epsilon$ et $ba \leftrightarrow \epsilon$. En quoi peut se réduire le mot $aaabababbbb$ ? A quelle condition deux mots sont-ils équivalents ?
6. Soit le Système de Réécriture suivant sur l'alphabet {**a, b, c**}. L'axiome est **bc**. Les règles de réécriture sont :
   - R1 tout mot $\phi$ sur l'alphabet peut se réécrire **a**$\phi$**a**
   - R2 tout mot $\phi$ de la forme $\alpha$**b**$\alpha'$**c**$\alpha''$ peut se réécrire $\alpha$**b**$\alpha'$**ac**$\alpha''$**a**

On appelle ”théorème” tout mot que l’on peut obtenir par dérivation à partir de l’axiome.
- Donner dans ce système la dérivation des ”théorèmes” **aaabcaaa, aaabaacaaaaa, baacaa**
- Les mots suivants sont-ils des ”théorèmes” :
– **aabaaacaa**
– **abca**
– **abcaaa**
– **abaca**
– **abacaa**
- Trouver un critère permettant de deviner à l’avance (i-e avant d’essayer de faire la dérivation) si un mot donné est un théorème ou non.
- Peut-on trouver une grammaire hors-contexte pour ce système ?

7. **Grammaires**
   Soit G la grammaire définie par $V_N = \{S, A, B\}$ et $V_T = \{a, b\}$ et :

$$\begin{array}{lllllll} S & \rightarrow & aB & & A & \rightarrow & a \\ S & \rightarrow & bA & & B & \rightarrow & bS \\ A & \rightarrow & aS & & B & \rightarrow & aBB \\ A & \rightarrow & bAA & & B & \rightarrow b & \end{array}$$

   Quel langage engendre-t-elle ?

8. Soit G la grammaire suivante :
   $V_N = \{SDP, TVP, NP, VP\}$,
   $V_T = \{child,\ doctor,\ every,\ some,\ no,\ Bill,\ Mary,\ John,\ laughed,\ cried,\ hugged,\ criticized\}$,
   avec les règles :

$$\begin{array}{lllllll} N & \rightarrow & child \mid doctor & & NP & \rightarrow & Bill,\ Mary,\ John \\ DP & \rightarrow & every \mid some\ \ no & & VP & \rightarrow & laughed \mid cried \\ TVP & \rightarrow & hugged \mid criticized & & & & \\ DP & \rightarrow & NP\,'s & & S & \rightarrow & NP\,VP \\ VP & \rightarrow & TVP\,NP & & NP & \rightarrow & DP\,N \end{array}$$

   + pour tout $C \in \{S, NP, VP, TVP\}$ :

$$\begin{array}{lll} C & \rightarrow & C \text{ and } C \\ C & \rightarrow & C \text{ or } C \\ C & \rightarrow & \text{neither } C \text{ nor } C \end{array}$$

   dériver dans cette grammaire :

- *Bill's child*
- *every doctor's child criticized Bill*
- *some doctor laughed and no child cried*
- *neither Bill nor John criticized Mary's doctor*
- *Bill and some doctor hugged and criticized Mary*

9. Soit G une grammaire où $V_N = \{S', S, N'', V'', V'', V', P'', prep, aux, comp, det\}$, l'axiome est $S'$, et où les règles sont :

$$\begin{array}{lcl} S' & \rightarrow & comp\ S \\ S & \rightarrow & N''\ V'' \\ V'' & \rightarrow & (aux)\ V' \\ V' & \rightarrow & V(N'')(P'')(S')(prep\ V'') \\ N'' & \rightarrow & (det)\ N' \\ N' & \rightarrow & N\ (S') \\ P'' & \rightarrow & prep\ N'' \end{array}$$

   Les parenthèses signifient l'optionnalité, par exemple $(X)$ signifie : X est possible mais pas nécessaire. On admet que $comp$ peut se réécrire $\emptyset$.
   - Montrer qu'avec un lexique approprié, cette grammaire permet d'engendrer les phrases :
     *Pierre dit que les journalistes mentent*
     *Pierre prédit à sa soeur qu'elle va réussir son examen*
     *Pierre donne le conseil à sa soeur qu'elle devrait travailler*
     *Pierre dit à son frère de rapporter du pain*
   - Dessiner les arbres de dérivation
   - dans la phrase *Pierre prédit à sa soeur qu'elle va réussir son examen*, les parties suivantes sont-elles des constituants :
     - *à sa soeur*
     - *prédit à sa soeur*
     - *sa soeur qu'elle va réussir son examen*
     - *réussir son examen*

10. Soit G la grammaire définie par $V_N = \{A, B, C\}$ et $V_T = \{0, 1\}$ et :

$$\begin{array}{lcl} A & \rightarrow & 0 \mid A0 \mid B1 \\ B & \rightarrow & 1 \mid A1 \mid C0 \\ C & \rightarrow & B0 \mid C1 \end{array}$$

   Démontrer que tout mot de $V_T^* - \{\epsilon\}$ est un A, un B ou un C.
   Si A est l'axiome, donner quelques exemples de mots de ce langage. Qu'ont-ils de particulier ?

11. Soit G la grammaire d'axiome S telle que $V_T = \{a, b, c\}$, $V_N = \{S, A, B, C, D\}$, avec les règles :

$$\begin{array}{lcl} S & \rightarrow & AB \mid CD \\ A & \rightarrow & aAb \mid ab \\ B & \rightarrow & Bc \mid c \\ C & \rightarrow & aC \mid a \\ D & \rightarrow & bDc \mid bc \end{array}$$

Montrer l'ambiguïté du mot $aabbcc$.

12. Soit G la grammaire donnée par : $V_N = \{S\}$ et $V_T = \{a\}$ et les règles :

$$\begin{array}{lcl} S & \rightarrow & aS \\ S & \rightarrow & aaS \\ S & \rightarrow & a \\ S & \rightarrow & aa \end{array}$$

Combien y a-t-il d'arbres syntaxiques différents pour engendrer le mot $aaaaa$ ?