

**Cursus:** M1, computer-science  
**Code UE:** JEIN8602  
**Subject:** Formal languages theory  
**Date:** 26 August 2014  
**Duration:** 3H  
 Documents: authorized  
 Lectures by: Mr Géraud Sénizergues

The exercises are independant one from each other.  
 It is *not* required to solve *all* the exercises. Every correct solution to an exercise will give (around) 4 marks.

**Exercice 1** [/4] Let us consider the regular expression:

$$e := ((abb)^* a(cb)^*) \cup (ab)^* (ba)^*$$

Construct, by Glushkov's method, a finite automaton recognizing  $L_e$ .

**Exercice 2** [/4] We consider the finite automaton  $\mathcal{A}$  described on figure 1. Note that 0 is

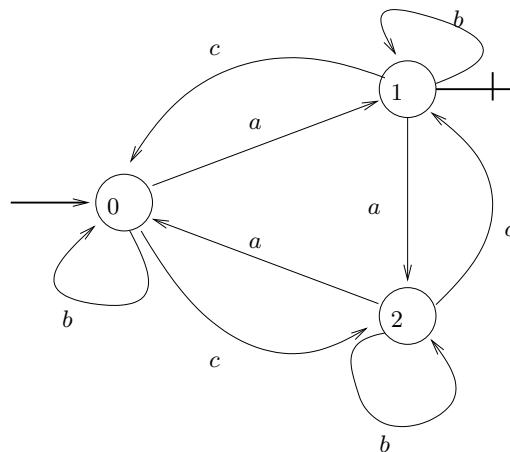


Figure 1: finite automaton  $\mathcal{A}$

the only initial state and 1 is the only final state.  
 Construct a regular expression for the language  $L_{\mathcal{A}}$  recognized by the automaton  $\mathcal{A}$ . Explain the successive steps of your construction.

**Exercice 3** [/5] Let

$$e = (abc)^*$$

Give a regular expression for the language  $\{a, b, c\}^* - L_e$  (i.e. the complement, in the set  $\{a, b, c\}^*$ , of the language defined by  $e$ ).

**Exercise 4** [/4] We consider the context-free grammar  $G := (A, N, R)$  where  $A = \{a, b\}$ ,  $N = \{S, T_1, T_2\}$  and  $R$  consists of the following rules:

**r1:**  $S \rightarrow T_1$

**r2:**  $S \rightarrow T_2$

**r3:**  $T_1 \rightarrow aT_1$

**r4:**  $T_1 \rightarrow aT_1b$

**r5:**  $T_1 \rightarrow a$

**r6:**  $T_2 \rightarrow T_2b$

**r7:**  $T_2 \rightarrow aT_2b$

**r8:**  $T_2 \rightarrow b$

The start symbol of  $G$  is  $S$ .

1- Which one of the following sets is exactly equal to  $L(G, T_1)$ ?

$$\{a^p b^q \mid p \geq q \geq 1\}, \quad \{a^p b^q \mid p > q \geq 1\}, \quad \{a^p b^q \mid p \geq q \geq 0\}, \quad \{a^p b^q \mid p > q \geq 0\},$$

2- Describe the language  $L(G, T_2)$  by an expression analogous to the expressions given at question 1.

3- Describe the language  $L(G, S)$  by an expression analogous to the expressions given at question 1.

4- Is the grammar  $G$  ambiguous ?

Indication: consider the word  $aaab$ .

5- What is the set of words which have exactly one derivation tree from  $S$  within the grammar  $G$  ?

6- Construct a non-ambiguous context-free grammar  $G'$ , with axiom  $S'$ , such that  $L(G, S) = L(G', S')$ .

**Exercise 5** [/4] We consider the context-free grammar  $G := (A, N, R)$  where  $A = \{a, b, c\}$ ,  $N = \{S_1, S_2, S_3, S_4, S_5\}$  and  $R$  consists of the following 12 rules:

$$\begin{array}{lll} S_1 \rightarrow aS_1S_1 & S_1 \rightarrow bS_3S_1 & S_1 \rightarrow S_2c \\ S_2 \rightarrow S_2S_1 & S_2 \rightarrow aS_3 & S_2 \rightarrow S_1S_2S_1 \\ S_3 \rightarrow a & S_3 \rightarrow S_1S_3 & S_4 \rightarrow cS_4 \\ S_4 \rightarrow aS_4S_5 & S_5 \rightarrow aS_5 & S_5 \rightarrow aS_4a \end{array}$$

The start symbol of  $G$  is  $S_1$ .

1- What are the *productive* non-terminals of  $G$  ?

2- What are the *useful* non-terminals of  $G$  ?

3- Transform the grammar  $G$  into an equivalent grammar  $G'$  where every non-terminal is productive and useful.

4- Is the language  $L(G, S_1)$  empty ?

5- Is the language  $L(G, S_1)$  infinite ?

**Exercise 6** [/4] We consider the two following context-free grammars  $G_1 := (A, N_1, R_1)$ ,  $G_2 := (A, N_2, R_2)$  where  $A = \{a, b, c\}$ ,  $N_1 = \{S, T\}$ ,  $N_2 = \{U\}$ ,  $R_1$  consists of the rules:

**r1:**  $S \rightarrow aSbT$

**r2:**  $S \rightarrow cT$

**r3:**  $T \rightarrow aTTb$

**r4:**  $T \rightarrow c$

and  $R_2$  consists of the rules:

**r4:**  $U \rightarrow UUb$

**r5:**  $U \rightarrow a$

- 1- Construct a context-free grammar generating the language  $L(G_1, S) \cdot L(G_2, U)$ .
- 2- Construct a context-free grammar generating the language  $L(G_1, S) \cup L(G_2, U)$ .
- 3- Construct a context-free grammar generating the language  $L(G_1, S)^*$ .
- 4- Let  $\varphi : \{a, b, c\}^* \rightarrow \{x, y\}^*$  be the homomorphism defined by:

$$\varphi(a) = xy, \quad \varphi(b) = yx, \quad \varphi(c) = y$$

We recall this means that, for every word  $w = a_1 \cdots a_i \cdots a_n$ , where the letters  $a_1, \dots, a_i, \dots, a_n$  belong to  $\{a, b, c\}^*$ :

$$\varphi(w) = \varphi(a_1) \cdots \varphi(a_i) \cdots \varphi(a_n)$$

and  $\varphi(\varepsilon) = \varepsilon$ . For example:

$$\varphi(ab) = xy yx, \quad \varphi(baa) = yxxyxy, \quad \varphi(cba) = yyxyxy.$$

Construct a context-free grammar  $H_1$  generating  $\varphi(L(G_1, S))$  and a context-free grammar  $H_2$  generating  $\varphi(L(G_2, U))$ .