TD4 - Formal grammars

March 14, 2025

1. [*Homework 23/24*]

   (a) Let G be the grammar $S \to aSbb \mid \varepsilon$. Describe informally the language it generates.

   (b) Let G' be the grammar $S' \to SSS, S \to aSbb \mid \varepsilon$, with S' the start symbol (axiom). Describe informally the language it generates.

   (c) Let G_1 and G_2 be context-free grammars, and $L(G_1)$ and $L(G_2)$ the languages they generate. How do you form context-free grammars generating the following sets:

      i. $L(G_1) \cup L(G_2)$
      ii. $L(G_1)L(G_2)$
      iii. $L(G_1)^*$

2. [*Exam 22/23*] Propose a (context-free) grammar that engenders the language ab^ncb^na with $n \geq 0$.

3. Consider the following grammars. For each grammar, give the engendered language and characterise it according to the Chomsky hierarchy.

$$\begin{array}{lcl} S & \to & S_1S_2 \\ S_1 & \to & aS_1b \mid ab \\ S_2 & \to & cS_2 \mid c \end{array} \qquad \begin{array}{lcl} S & \to & aSBC \\ S & \to & aBC \\ CB & \to & BC \\ aB & \to & ab \\ bB & \to & bb \\ bC & \to & bc \\ cC & \to & cc \end{array}$$

4. [*Homework 21/22*] A noun phrase is either a determiner followed by a noun or a noun phrase followed by a prepositional phrase. A prepositional phrase is a preposition followed by a noun phrase.

   (a) Write the grammar G of noun phrases.

   (b) Write the two derivation trees that G associates with the word `D N P D N P D N`, where `D` is the symbol for the category 'determiner', `N` for 'noun' and `P` for 'preposition'.

   (c) Which of the two trees of question 4b would you choose to represent the structure of the NP 'le chat de la voisine de la concierge'?

   (d) Propose a grammar G', different from G, that associates to `D N P D N P D N` only the analysis that you have identified in the previous question.

   (e) Assuming that `D`, `N` and `P` are terminal symbols, propose a regular grammar equivalent to G'.

5. [*Exam 22/23*] Let us consider the language L_1 on the vocabulary $V = \{$*the*, *man*, *bear*, *that*, *saw*$\}$ that include all finite sentences of the form (*the man that saw*)$^+$ *the bear*.

   (a) Give a context-free grammar that generates L_1.

   Let L_2 be the language engendered by the grammar G_2:

$$\begin{array}{lcl} \text{S} & \to & \text{NP Rel} \\ \text{NP} & \to & \textit{the man} \\ & \mid & \textit{the bear} \\ \text{Rel} & \to & \textit{that saw} \text{ NP Rel} \\ & \mid & \textit{that saw the bear} \end{array}$$

   (b) Propose a grammar G_3 that engenders the language L_3, which is a superset of L_2 in which the symbols *bear* and *man* are **interchangeable**.

   (c) Describe informally the differences between L_1 and L_2.

   (d) Give a context-free grammar of the language $L_2 \setminus L_1$.