# Homework 3 Automata Theory 2023

Published on: Wednesday 15 November 2023.
Deadline for submission: Wednesday 29 November 2023, 23:59.
The assignment must be completed individually. A total of 100 points can be earned. Answers to be submitted via Brightspace. Submit a single file, e.g., a pdf or possibly a zip. Please include your name and student number in your submission. You may either type your answers or hand-write them. In the latter case, please hand in an easy-to-read scan / photos.

1. [25 pt] Let $L$ be the language consisting of all strings $x \in \{a, b\}^*$, such that

   - $n_b(x) \geq 1$, and
   - after the last occurrence of $b$, $x$ contains at least $n_b(x)$ $a$'s, and
   - $n_a(x) > n_b(x)$, i.e., in addition to the $a$'s from the previous condition, $x$ contains at least one more $a$ (at some point in the string).

   Hence, the first five elements in the canonical (shortlex) order of $L$ are: $aba, baa, aaba, abaa, baaa, aaaba$. But also, e.g., $abbaa$ and $babaa$ are elements of $L$.

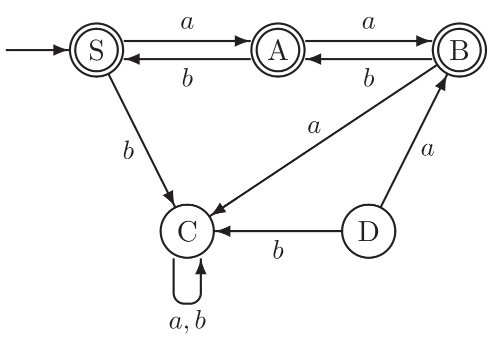
   (a) Give a context-free grammar $G$, such that $L(G) = L$.

   (b) Explain why $G$ indeed generates exactly $L$ (you do not need to prove this formally).

2. [30 pt] In a regular grammar $(V, \Sigma, S, P)$, every production is of one of the following forms:

   - $A \to \sigma B$, with $A, B \in V$ and $\sigma \in \Sigma$
   - $A \to \Lambda$, with $A \in V$

   Now consider the following finite automaton $M$:




   (a) Use the construction from Theorem 4.14 (in Section 4.3 of the book) to construct a regular grammar $G$ generating $L(M)$. Give as your answer the resulting grammar $G$.

   (b) Determine step by step (hence via $N_0, N_1, N_2, \ldots$) the live variables in your grammar $G$.
   Next, give the grammar $G'$ resulting from $G$ by removing the variables that are *not* live.

   (c) Determine step by step (hence via $N_0, N_1, N_2, \ldots$) the reachable variables in $G'$.
   Next, give the grammar $G''$ resulting from $G'$ by removing the variables that are *not* reachable.

3. [25 pt] Let $G$ be the context-free grammar with start variable (and only variable) $S$, and the following productions:

$$S \to SaS \ | \ b \ | \ \Lambda$$

Let $x$ be an arbitrary string in $\{a, b\}^*$, that has no occurrence of the substring $bb$. Use induction on the length $|x|$ of $x$ to prove that $x \in L(G)$, i.e., that $x$ can be generated by $G$.

*Hint:* For $|x| \geq 2$, use that $x$ cannot consist of all $b$'s.

4. [20 pt] Let $G$ be the context-free grammar with start variable $S$ and the following productions:

$$S \to aAbB \mid b \qquad A \to BBa \mid bb \qquad B \to bS \mid AB$$

This grammar does not contain $\Lambda$-productions or unit productions.

(a) Give the context-free grammar $G'$ resulting from $G$ by introducing for every terminal symbol $\sigma$ a variable $X_\sigma$ (with a corresponding production), and substituting this variable for occurrences of $\sigma$ where necessary in the righthand side of productions.

(b) Give the context-free grammar $G''$ resulting from $G'$ by splitting the righthand side of productions which are too long.