Examination,
Finite automata and formal languages (DIT322/TMV028)

- Date and time: 2021-08-18, 8:30–12:30.
- Examiner: Nils Anders Danielsson.
- The GU grades Pass (G) and Pass with Distinction (VG) correspond to the Chalmers grades 3 and 5, respectively.
- To get grade n on the exam you have to be awarded grade n or higher on at least n exercises.
- A completely correct solution of one exercise is awarded the grade 5. Solutions with minor mistakes *might* get the grade 5, and solutions with larger mistakes might get lower grades.
- Exercises can contain parts and/or requirements that are only required for a certain grade. To get grade n on such an exercise you have to get grade n or higher on every part marked with grade n or lower (and every unmarked part), and you have to fulfil every requirement marked with grade n or lower (as well as every unmarked requirement).
- Answers can be written in Swedish or English.
- Answers must be given in files with one of the following formats: PDF, JPEG or TXT. Submit your solutions to Canvas before the deadline. Note that there is a separate Canvas assignment for each of the six questions. If Canvas is not working properly, send the solutions to the examiner using email, and include the course code in the subject header.
- Solutions can be rejected if they are hard to read (for instance if an image is out of focus), unstructured, or poorly motivated.
- You do not need to provide proofs showing that algorithms covered in the course are correct. It is fine to use arguments of the following form: "Here I have used algorithm X to compute the value y, and because the result of algorithm X always satisfies property P, we have $P(y)$." However, you have to explain step by step why algorithm X produces the value y.
- **No collaboration is permitted, you have to work on your own.**
- If you want to discuss the grading of the exam, contact the examiner no later than three weeks after the result has been reported.

1. Consider the context-free grammar $G = (\{ A, B, S \}, \{ a, b \}, P, S)$, where the set of productions P is defined in the following way:

$$S \to AB \qquad A \to a \mid AA \qquad B \to b \mid BB$$

   (a) Is G in Chomsky normal form (as defined in the lectures)?

   (b) If G is in Chomsky normal form, construct the CYK table for G and the string abb. No proof or explanation is required.

   (c) Construct a parse tree (in $P(G, S)$) with yield aba, or prove that there is no such parse tree.

2. Give a pushdown automaton A such that

$$N(A) = \{ 0^n 1^n \mid n \in \mathbb{N} \}.$$

   The automaton's input alphabet should be $\{ 0, 1 \}$.

   Explain your answer in such a way that the person correcting your exam can easily see that the answer is correct.

3. Which of the following regular expressions, if any, denote the same language?

$$e_1 = (1^*)^*$$
$$e_2 = (1^*)^+$$
$$e_3 = (1^+)^*$$

   Provide proofs of both equalities and inequalities. You are allowed to make use of the following laws without providing proofs for them:

   - Laws presented in the lectures.
   - $L(e^*) = L(\varepsilon + ee^*)$.
   - $L((e^*)^*) = L(e^*)$.
   - $L(e^*e^*) = L(e^*)$.
   - If $L(e_1) \subseteq L(e_1')$ and $L(e_2) \subseteq L(e_2')$ then $L(e_1^*) \subseteq L(e_1'^*)$, $L(e_1 e_2) \subseteq L(e_1' e_2')$ and $L(e_1 + e_2) \subseteq L(e_1' + e_2')$.

   You are also allowed to make use of algorithms presented in the lectures.

4. Consider the regular expression $e = ((ab)^* + b)^*$.

   (a) Construct an ε-NFA A over the alphabet $\{ a, b \}$ satisfying $L(A) = L(e)$.

   Prove that your construction is correct.

   (b) Construct a minimal DFA B over the alphabet $\{ a, b \}$ satisfying $L(B) = L(e)$.

   Prove that your construction is correct.

   (c) Give a precise description, using natural language, of the language $L(e)$. Explain your answer in such a way that the person correcting your exam can easily see that the description is correct.

5. Let the context-free grammar G be $(\{ S, A \}, \{ a, b \}, P, S)$, where P contains exactly the following productions:

$$S \to SaA \mid b \qquad A \to Sa$$

   (a) *For grade 3:* Define a language L as an inductively defined subset of $\{ a, b \}^*$ (by giving some inference rules) in such a way that $L = L(G)$. It is OK to simultaneously define one or more other languages (also as inductively defined subsets of $\{ a, b \}^*$).

   Explain your answer in such a way that the person correcting your exam can easily see that the answer is correct. (The explanation could refer to parts (b) and (c) below.)

   (b) *For grade 4:* Prove that $L \subseteq L(G)$.

   (c) *For grade 4:* Prove that $L(G) \subseteq L$.

   For grade four it suffices to complete one of the proofs. For grade five both proofs are required.

6. Consider the languages M and N over the alphabet $\{ a, b, c \}$, where

$$M = \left\{ awcwb \;\middle|\; w \in \{ a, b \}^* \right\}$$

   and N is given as an inductively defined subset of $\{ a, b, c \}^*$:

$$\frac{}{b \in N} \qquad \frac{u \in \{ a, b \}^*}{u \in N} \qquad \frac{u, v \in N}{ucv \in N}$$

   For each language, answer the following two questions:

   - Is the language regular?
   - Is the language context-free?

   Provide proofs. You are allowed to make use of the following lemmas/facts without proving them:

   - The two pumping lemmas covered in the course.
   - The closure properties covered in the course.
   - The fact that for any alphabet Σ with at least two elements the language $\{ ww \mid w \in \Sigma^* \}$ is not context-free.
   - The fact that for any alphabet Σ with at least two elements the language $\left\{ ww^{\mathrm{R}} \;\middle|\; w \in \Sigma^* \right\}$ is context-free and not regular.
   - The fact that if Σ is an alphabet and $L \subseteq \Sigma^*$ is context-free, then, for any string $w \in \Sigma^*$, $\{ v \in \Sigma^* \mid wv \in L \}$ and $\{ v \in \Sigma^* \mid vw \in L \}$ are also context-free (see "Quotients of Context-Free Languages" by Ginsburg and Spanier).