Aalto University, School of Science
Department of Computer Science
Sándor Kisfaludi-Bak

CS-C2160 Theory of Computation (5 cr)
Course Exam 19 April 2024, 09:00–12:00 p.m.

Instructions

You are allowed to bring with you a single two-sided A4 "cheat sheet", **personally handwritten by you. (NO photocopies, NO printouts, NO computer type-set text.)** Please include your name and student ID at the top of the cheat sheet, and return it together with your answer sheets at the end of the exam.

Note: If you have not completed your online A+ home assignments, your exam will not be graded.

Problems

1. Design **deterministic** finite automata for recognising the following languages:

   (a) $\{w \in \{a,b\}^* \mid w$ contains the substring abb (as consecutive symbols)$\}$,

   (b) $\{w \in \{a,b,c\}^* \mid w$ has at most 3 symbols that are not $c\}$,

   (c) $\{w \in \{a,b,c\}^* \mid acab$ appears as a subsequence of w (not necessarily consecutively)$\}$.

   Give **regular expressions** that describe the following languages:

   (d) $\{w = a_1a_2,\dots,a_{2n} \in \{0,1\}^* \mid n \in \mathbb{N}$, and (all 1 symbols of w are in even positions or all 1 symbols of w are in odd positions)$\}$

   (e) $\{w \in \{0,1\}^* \mid$ **if** w ends with 00 **then** it starts with $11\}$. 15 points

2. For a string $w \in \{a,b,c\}^*$, let us denote by $n_s(w)$ the number of s symbols in w. Design context-free grammars for the following languages:

   (a) $L_{3k+1} = \{w \in \{a,b\}^* \mid n_a(w) = 3k+1$ for some $k \in \mathbb{N}\}$, (3pts)

   (b) $L_+ = \{w \in \{a,b,c\}^* \mid n_a(w) + n_b(w) = n_c(w)\}$, (3pts)

   Hints: In part (a), it is a good idea to first design a grammar for $L_{3k} = \{w \in \{a,b\}^* \mid n_a(w) = 3k$ for some $k \in \mathbb{N}\}$. In part (b), the symbols a and b play the same role.

   Give parse trees in your grammars for the following strings:

   (c) $aababa \in L_{3k+1}$, (2pts)

   (d) $cabacc \in L_+$. (2pts)

   Finally,

   (e) Convert the grammar $S \to aSbS|\varepsilon$ into Chomsky Normal Form. (5pts) 15 points

3. The symmetric difference of the languages A and B is the language

$$A \Delta B := (A \setminus B) \cup (B \setminus A).$$

(a) Let $A = \{0^k1^\ell0^k \mid k,\ell \in \mathbb{N}\}$ and $B = \{0^k1^\ell0^k \mid k,\ell \in \mathbb{N} \text{ and } k \neq \ell\}$. Describe the language $A\Delta B$ as simply as you can, without using the Δ operation, and prove that $A\Delta B$ is not a regular language.

(b) Let $M_{A'}$ and $M_{B'}$ be total Turing machines (deciders) for languages A' and B', respectively. Using these machines as building blocks, give a diagram or describe a total Turing machine (decider) for $A'\Delta B'$. (There are no restrictions on your Turing machine in terms of tape count or determinism, and you may use any Turing machines discussed in the lectures – e.g., the one that can clear a tape – as building blocks.)

(c) Prove that there are semi-decidable languages A'' and B'' where $A''\Delta B''$ is not semi-decidable.
Hint: It is best to give an explicit example. You can even choose one of the languages to be very "simple".

15 points

4. Which of the following claims are true and which are false? Provide a brief justification for each of your answers, based on results introduced at the course. Note that just stating "True" or "False" without justification will not earn you any points. (For example if the claim was: "The complement of any decidable language is semidecidable", your answer could be: "True. The complement of any decidable language is decidable (by switching the accepting and rejecting states in the recognising TM), and all decidable languages are by definition also semidecidable.")

   (a) Non-deterministic finite automata can recognize some languages that cannot be described with regular expressions.
   (b) The complement of any finite language is context-free.
   (c) The union of any regular language and any decidable language is context-free.
   (d) There are decidable languages A, C and an undecidable language B such that $A \subseteq B \subseteq C$.
   (e) There is an undecidable language B that is a subset of the regular language $L = \{w \in \{0,1\}^* \mid w \text{ starts with } 1001\}$.

15 points