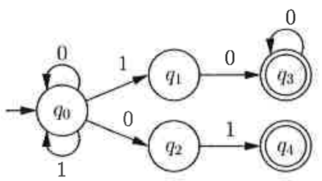
Aalto University, School of Science
Department of Computer Science
Viorel preoteasa (040-5609056)

CS-C2150 Theoretical Computer Science — Spring 2017

Exam April 5th, 2017, at 17:00–20:00

Use of calculators is not allowed in the exam.
Note: if you have not completed your computerized home assignments, your exam will not be graded.

1. (a) Design a deterministic finite state automaton that recognizes the language
$$\{w \in \{a,b\}^* \mid w \text{ contains the substring } baa\}.$$
   (b) Design a deterministic finite state automaton that recognizes the language
$$\{w \in \{a,b\}^* \mid w \text{ does not contain the substring } baa\}.$$
   (c) Let us consider the following finite non-deterministic automaton over the alphabet $\{0,1\}$:



   Describe the language recognized by the automaton verbally, in one or two sentences. Give a deterministic finite state automaton that recognizes the same language.

10 points

2. (a) Give a regular expression that describes the language
$$L = \{w \in \{a,b\}^* \mid w \text{ begins and ends with different symbols}\}$$
   (b) Consider the regular expression $((a^*abb) \cup (bb))^*$ over the alphabet $\{a,b\}$. Give the deterministic finite state machine with *minimal number of states* that recognizes the language described by the regular expression.
   (c) Give a regular expression that describes the language
$$L = \{w \in \{a,b,c\}^* \mid w \text{ does not contain the substring } aba\}$$
   Hint: it may be a good idea to first build a deterministic finite state automaton recognizing the language and then build the expression from the automaton.

10 points

3. Consider the language
$$L = \{ucv \mid u, v \in \{a,b\}^* \text{ and } |u| \leq |v|\}$$
   over the alphabet $\{a,b,c\}$.
   (a) Prove that the language is not regular.
   (b) Design a contex-free grammar that generates the language.
   (c) Give parse trees for the strings $aacaa$ and $babcbaba$ in your grammar.
   (d) Design a pushdown automaton that recognizes the language and describe its operation (design idea) with few sentences. Is your machine deterministic?

16 points

4. Consider a simplified context free grammar for a programming language in which the set of terminal symbols is $\Sigma = \{\textbf{if}, \textbf{then}, \textbf{while}, \textbf{do}, ;, :=, =, <, +, x, y, z, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$, the set of variables is $\{stmts, stmt, assign, cond, expr, num, digit, var\}$, *stmts* is the start variable, and the rules are

$$\begin{aligned}
stmts &\rightarrow stmt \mid stmt\ ;\ stmts \\
stmt &\rightarrow \textbf{if}\ cond\ \textbf{then}\ stmt \mid \textbf{while}\ cond\ \textbf{do}\ stmts \mid assign \\
assign &\rightarrow var := expr \\
cond &\rightarrow expr < expr \mid expr = expr \\
expr &\rightarrow var \mid num \mid expr + expr \\
num &\rightarrow digit \mid digit\ num \\
digit &\rightarrow 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9 \\
var &\rightarrow x \mid y \mid z
\end{aligned}$$

As an example, the Σ-string "$x := x + y + 3; \textbf{while}\ x < 2\ \textbf{do}\ x := x + 1$" is produced by the grammar.

(a) Prove that the grammar is ambiquous.

(b) Is the grammar left recursive? Justify your answer.

(c) Is the grammar an LL(1) grammar? Justify your answer briefly (you don't have to construct the full FIRST and FOLLOW sets).

6 points

5. (a) Desribe in your own words (with at most 5 sentences), what "Church–Turing thesis" is.

(b) Define the concepts "Turing-recognizable language" and "Turing-decidable language" ("recursively enumerable language" and "recursive language" in "Orposen pruju").

(c) Consider the language

$$L_{\text{composite}} = \{x \in \{0,1,2,3,4,5,6,7,8,9\}^* \mid x = y \times z \text{ for some integers } 1 < y < z\}.$$

For instance, the string 36 is in the language as $36 = 4 \times 9$ but 23 is not in the language. Is the language Turing-decidable? Justify your answer.

(d) Given a language L over an alphabet $\{0,1\}$, let $L^{\text{next}} = \{w^{\text{next}} \mid w \in L\}$ be the language obtained by computing the next element w^{next} of w in the canonical order $0 < 1 < 00 < 01 < 10 < 11 < 000 < \ldots$. For example $01011^{\text{next}} = 01100$, $111^{\text{next}} = 0000$. Show the following claim either correct or incorrect:

L is a Turing-decidable language if and only if L^{next} is a Turing-decidable language.

10 points

6. Consider the following decision problem:

Given a Turing machine M. Do all strings accepted by the machine M contain the symbol a at most 4 times?

It can be described as the language

$$L = \{\langle M\rangle \mid M \text{ is a Turing machine and for all } x \in L(M) \text{ the number of occurences of symbol } a \text{ in } x \text{ is less or equal to } 4\}.$$

Prove that the language L is undecidable. If you use Rice's theorem (you don't have to; you can, for instance, also use reductions), then also state Rice's theorem and the related concepts of "semantic property", "trivial semantic property", and "decidable semantic property".

7 points

7. At what time did you finish answering the exam questions?

1 points