RESIT AUTOMATA THEORY
Thursday 3 February 2022, 10.15 - 13.15

This exam consists of eight exercises, where [x pt] indicates how many points can be earned per exercise. A total of 100 points can be earned.
It is important to provide an explanation or motivation when a question asks for it.
A finite automaton in this exam (without further addition), refers to a deterministic finite automaton without Λ-transitions (which is elsewhere called *DFA*).

---

1. [9 pt] Let

$$L = \{x \in \{a,b\}^* \mid x \text{ contains at least an occurrence of } aa \text{ and an occurrence of } abb\}$$

Note that substring occurrences may overlap. For example, $aabb \in L$.

Draw a finite automaton M, such that $L(M) = L$.

---

2. [12 pt]

(a) Let $M_1 = (Q_1, \Sigma, q_1, A_1, \delta_1)$ and $M_2 = (Q_2, \Sigma, q_2, A_2, \delta_2)$ be any two finite automata with the same input alphabet Σ. Let $L_1 = L(M_1)$ and $L_2 = L(M_2)$. Using the product construction you can construct a finite automaton $M = (Q, \Sigma, q_0, A, \delta)$ from M_1 and M_2, such that $L(M) = L_1 \cup L_2$.

Describe (in words and/or formulas, but in any case clearly and completely) what are the components Q, q_0, A and δ of this automaton M.

(b) Explain how to modify the product construction from part (a) to accept the language:

$$\{x \in \Sigma^* \mid \text{if } x \in L_1, \text{ then also } x \in L_2\}$$

Also briefly explain why the modified construction is correct.

---

3. [16 pt]

   (a) Let $L \subseteq \Sigma^*$ be a language, and let $x \in \Sigma^*$ be a string.
   How is the set L/x defined (the 'future set' of x with respect to L)? That is, what strings are in L/x?

   If you do not know the answer to this question, you can 'buy' it from the lecturer. Perhaps you can then solve (b) and (c).

Let

$$L_1 = \{x \in \{a,b\}^* \mid x = a^i(ba)^j a^k \text{ with } i,j,k \geq 0$$
$$\text{and } x \text{ contains as many occurrences of the substring } ba \text{ as of the substring } aa\}$$

A string $x \in L_1$ must thus satisfy all stated requirements. Note that occurrences may overlap. For example, $x = aaa$ contains two occurrences of the substring aa.

Indeed, we know language L_1 from the homework assignments. The first eight elements in canonical (shortlex) order of L_1 are: $\Lambda, a, baa, aaba, abaa, babaaa, aaababa, aababaa$.

   (b) Let $m \geq 0$ be arbitrary,[1] and let $x = a^m$. Given the above language L_1, what is the set L_1/x? Distinguish in your answer the case $m = 0$ and the case $m \geq 1$.

   Describe in your answer concretely what are the elements of L_1/x , in terms of a, ba, m, i, j and k. Only if the 'future set' were the entire language L_1, you can simply say L_1.

   (c) In a previous exam we have seen that two strings a^m and a^n with $m, n \geq 0$ and $m \neq n$ are always L_1-distinguishable. In other words: that they have different future sets.

   Give two different, concrete strings $x, y \in \{a,b\}^*$, such that $L_1/x = L_1/y \neq \emptyset$. That is, two different strings that have the same, non-empty 'future set'. Give also the common set $L_1/x = L_1/y$. Describe in your answer concretely what are the elements of $L_1/x = L_1/y$, in terms of a, ba, i, j and k.

   N.B.: Obviously, at least one of the strings x and y must contain (at least) an occurrence of the letter b.

---

[1] This does not mean that you can *choose* any $m \geq 0$, but that m can be any non-negative number.

4. [10 pt] Let

$$L = \{x \in \{a,b\}^* \mid |x| \geq 2 \text{ and } x \text{ begins and ends with } aa \text{ or } ba\}$$

Some elements of L are, for example, aaa and $aababa$. For each of the following regular expressions r_1, r_2, r_3, r_4 indicate whether it is a correct expression for L or not. If an r_i is not a correct expression, explain this, by giving a string x

(a) that is in L, but does not satisfy r_i,

(b) or which is not in L, but satisfies r_i.

If strings x can be found for both situations, you only need to provide one string x. Clearly indicate whether situation (a) or (b) applies for your provided string x.

$$\begin{aligned} r_1 &: \quad (a+b)(a+b)^*a \\ r_2 &: \quad (a+b)a(a+b)^*a + aa \\ r_3 &: \quad (aa+ba)(b^*a)^* \\ r_4 &: \quad (aa+ba)((a+b)^*(aa+ba))^* \end{aligned}$$

---

5. [10 pt] Below are given two grammars G_1 and G_2, that were once proposed for the language L_1 from exercise 3. The grammars are not necessarily correct for language L_1, but that is not the point. For every G_i, answer the following questions:

Is G_i ambiguous?

If yes, show this by giving two different derivations trees for the same string $x \in L(G_i)$.

If not, show this by reasoning that different productions in G_i yield different strings. You do not need to provide a complete induction proof.

(a) G_1 is the context-free grammar with start variable S and the following productions:

$$\begin{aligned} S &\rightarrow AB \mid aAB \mid \Lambda \\ A &\rightarrow aAba \mid \Lambda \\ B &\rightarrow baBa \mid \Lambda \end{aligned}$$

(b) G_2 is the context-free grammar with start variable S and the following productions:

$$\begin{aligned} S &\rightarrow aA \mid B \mid aaAbabaBa \\ A &\rightarrow aAba \mid \Lambda \\ B &\rightarrow baBa \mid \Lambda \end{aligned}$$

---

6. [15 pt]

   Let

   $$L = \{a^i b^j c^k \mid i, j, k \geq 0 \text{ and } i < j + k\}$$

   (a) Give the first six elements in canonical (shortlex) order of L.

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

   Also, explain what is the role of the various variables and productions in G to generate L.

---

7. [14 pt]

   Let

   $$L = \{a^i b^j c^k \mid i, j, k \geq 0 \text{ and } i \neq j + k\}$$

   N.B.: this language is slightly different from the language L in exercise 6. Now i may also be greater than $j + k$.

   Draw a pushdown automaton M, such that $L(G) = L$. Try to ensure that M is deterministic and does not contain any Λ-transitions. If you do not succeed in this, you can still earn most of the points.

   Also explain how M uses its states and stack to accept precisely the right language.

---

8. [14 pt] Let G be the context-free grammar with start variable S and the following productions:

   $$S \to aSb \mid T \qquad T \to aT \mid \Lambda$$

   (a) Draw the non-deterministic top-down pushdown automaton $NT(G)$.

   If you do not know the answer to this question, you can 'buy' it from the lecturer. Perhaps you can then solve (b) and (c).

   (b) Perform a successful computation in $NT(G)$ for the input $x = aab$, i.e.: a computation that starts in the initial configuration for x and results in the acceptance of x.

   Give this computation as in the lectures with a table containing columns for (1) the state, (2) the already matched input, (3) the remaining input, (4) the stack contents, and (5) the expand- or match-move (if applicable).

   (c) Give the leftmost derivation of aab in G that corresponds to the computation in part (b).

`end of exam`