Middle East Technical University Department of Computer Engineering

CENG 280

Formal Languages and Abstract Machines

Spring 2018-2019

Take Home Exam 1

Due date: 19 March 2019, 23:59

Objectives

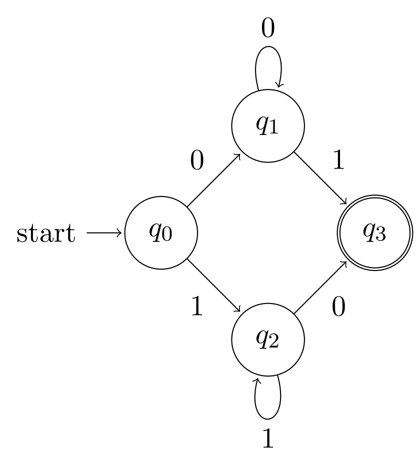
Reviewing essential concepts of algebra that are extensively used in formal languages theory, this homework aims to familiarize you with *finite automaton (FA)*, the most restricted abstract computing device to recognize a *Regular Language*, observing the benefits and limitations of employing such simplistic theoretical machinery to tackle practical problems and to help you gain experience in equivalent yet distinct ways of finite representation of the same class of languages such as *Regular Expressions*.

Specifications

You must adhere to the notation conventions adapted in the textbook. All automata must be defined as tuples and you must illustrate state transition functions and state transition relations graphically.

Your solution should be delivered as a .pdf file advisably generated using LaTeX. For convenience, a simple code for drawing an automaton is included in the following. On the left-hand side you can see the code segment, and generated automaton is placed on the right.

```
\begin{tikzpicture}[shorten >=1pt,node distance=2cm,on
    grid,auto]
\node[state,initial] (q_0) {$q_0$};
\node[state] (q_1) [above right=of q_0] {$q_1$};
\node[state] (q_2) [below right=of q_0] {$q_2$};
\node[state,accepting](q_3) [below right=of q_1] {$q_3$};
\path[->]
(q_0) edge node {0} (q_1)
edge node [swap] {1} (q_2)
(q_1) edge node {1} (q_3)
edge [loop above] node {0} ()
(q_2) edge node [swap] {0} (q_3)
edge [loop below] node {1} ();
\end{tikzpicture}
```



Questions and submission regulations are included in subsequent sections. Designing your solutions to the tasks, explicitly state any assumptions you make and pay particular attention to the notation you use. Your proofs must be sound and complete. Grading will be heavily affected by the formalization of your solutions.

Question 1 (12 pts)

The sets

$$A_k = \{w : w \in \{0,1\}^*, |w| \leq k\}$$
$$B = \mathcal{L}(0^*10(0 \cup 1)^*)$$
$$C = 2^{\{0,1\}^*}$$

where $k \in \mathbb{N}$ are given.

Construct formal proofs to state whether the following sets are **countable** or **uncountable**. If any of them is countable, you should also show whether they are **finite** or **infinite**, i.e. write down their cardinalities in these cases. State any mapping explicitly and give clear references to known theorems if used. Any other assertions must be proven by yourselves so that your concise final answers are not treated as random guesses.

a. $C \setminus \{\{x\} : x \in (A_{280} \cup B)\}$ (3 pts)

b. $A_7 \cap B^* \cap C$ (3 pts)

c. $\bigcup C \setminus (A_2 \times B)$ (3 pts)

d. $(\bigcup C \cup \bigcup_{k=1}^{\infty} A_k) \setminus \{0,1\}^*$ (3 pts)

(Note that $\bigcup S = \{x : x \in P \text{ for some set } P \in S\}$ and $\bigcup_{k=1}^{n} S_k = S_1 \cup S_2 \cup ... S_n$.)

Question 2 (13 pts)

All finite automata (FA) with **four pre-defined states** on a **binary alphabet** have definitions as five tuples $M = (K, \Sigma, \delta, s, F)$ such that $K = \{q_0, q_1, q_2, q_3\}$ and $\Sigma = \{0, 1\}$ are given, yet other constituents are **unfixed**, i.e. they can take on many distinct values. Assume that two FA are **not different** if and only if all of the associated five tuples are equal to each other. Pay attention to the distinction between the equivalence of sets and the equivalence of ordered pairs.

a. How many different **deterministic** FA M exist in accordance with given specifications? (5 pts)

b. How many different **nondeterministic** FA M exist in accordance with given specifications? (5 pts)

c. How do you explain the difference between options (**a**) and (**b**)? (1 pts)

d. How do the numbers obtained in options (**a**) and (**b**) relate to the number of different languages corresponding FA M recognize? Justify your answer in a few sentences. (2 pts)

Question 3 (15 pts)

For each language L_i, write a **regular expression** α_i representing the language, i.e. $L_i = \mathcal{L}(\alpha_i)$ s.t. $i \in \{1, 2, 3\}$. Interpret symbols 0 and 1 as natural numbers in related contexts.

a. $L_1 = \Big\{ w \in \{0,1\}^* : \Big(\sum_{i=1}^{|w|} w(i)\Big)$ is a multiple of 2 and 4, $w(i) + w(i+1) \neq 2$ for $i \in [1..(|w|-1)] \Big\}$. (5 pts)

b. $L_2 = \Big\{ w \in \{0,1\}^* : |w|$ is odd, and $w(2i) = 0$ for $i = 1, 2, ... \Big\lfloor \frac{|w|}{2} \Big\rfloor \Big\}$. (5 pts)

c. $L_3 = \{w \in \{0,1\}^* : w$ has even number of interleaved occurrences of the substring $00\}$. (5 pts)

(*Note*: $000 \notin L_3$ since it has one interleaved occurrence of 00, yet $0000 \in L_3$ and $00000 \in L_3$.)

Question 4 (14 pts)

In each part, construct a **deterministic finite automaton (DFA)** that recognizes the given language. **Draw** all your deterministic finite automata (DFA) using graphical notation and write down their constituents as five tuples in set/tuple notation, **leaving out** the state transition function which must only be depicted graphically. Explicitly draw transitions to trap (dead) states whenever applicable. Interpret symbols 0 and 1 as natural numbers in related contexts.

a. $L_1 = \Big\{ w \in \{0,1\}^* : \Big(\sum_{i=1}^{|w|} w(i)\Big) \geq 2$ and $\Big(|w| \; - \; \sum_{i=1}^{|w|} w(i)\Big) \leq 2 \Big\}$. (7 pts)

b. $L_2 = \{w \in \{0,1\}^* : w$ contains neither 1011 nor 00 but has 0101 as a substring$\}$. (7 pts)

Question 5 (8 pts)

Given the **nondeterministic finite automaton (NFA)** N=(K, $\{a, b\}$, Δ, q_0, F), in which $K = \{q_0, q_1, q_2, q_3, q_4\}$,

$$\begin{aligned}\Delta = \{&(q_0, b, q_1), (q_0, e, q_2), (q_0, b, q_3),\\ &(q_1, a, q_1),\\ &(q_2, a, q_2), (q_2, b, q_2), (q_2, a, q_4),\\ &(q_3, a, q_1), (q_3, e, q_4),\\ &(q_4, a, q_1), (q_4, a, q_2), (q_4, b, q_4)\}\end{aligned}$$

and $F = \{q_1, q_3\}$, trace the following strings on N employing the **configuration** notation for nondeterministic finite automata (NFA) within **yields-in-one-step** relation and decide whether they are in $L(N)$ or not utilizing the **reflexive transitive closure** of yields-in-one-step relation. Answers based on reasoning about the type of strings N accepts / rejects will be graded **zero**.

a. $w_1 = abaa$. (4 pts)

b. $w_2 = babb$. (4 pts)

Question 6 (9 pts)

Given the **NFA** N=(K, $\{a, b\}$, Δ, q_0, F), in which $K = \{q_0, q_1, q_2, q_3, q_4\}$,

$$\begin{aligned}\Delta = \{&(q_0, a, q_0), (q_0, b, q_0), (q_0, b, q_1), (q_0, e, q_2),\\ &(q_1, a, q_4),\\ &(q_2, b, q_1), (q_2, b, q_3),\\ &(q_3, a, q_4)\\ &(q_4, e, q_3)\}\end{aligned}$$

and $F = \{q_3, q_4\}$, construct an equivalent **DFA** M using the **subset construction algorithm** for NFA to DFA conversion in the textbook. Precisely show **every** step of the algorithm applied on N to yield M. Answers comprising of only the resultant DFA M will be graded **zero**.

Question 7 (9 pts)

Given the **NFA** N=(K, $\{a, b\}$, Δ, q_0, F), in which $K = \{q_0, q_1, q_2, q_3\}$,

$$\begin{aligned}\Delta = \{&(q_0, b, q_1), (q_0, a, q_3),\\ &(q_1, a, q_1), (q_1, a, q_3),\\ &(q_2, a, q_2), (q_2, b, q_2),\\ &(q_3, b, q_1), (q_3, b, q_2)\}\end{aligned}$$

and $F = \{q_2\}$, construct a regular expression α such that $\mathcal{L}(\alpha) = L(N)$ using the **FA to regular expression conversion method** in the textbook. Begin by writing down the **generalized finite automaton (GFA)** for N. Then, explicitly show **every** step of the method's application on the GFA. Simplify the obtained regular expression if possible. Answers that do not show interim computations will be graded **zero**.

Question 8 (10 pts)

Let L be a language over a finite alphabet Σ and assume that $\{0, 1\}$ is a subset of Σ. Define another language H as

$$H = \{w \in \Sigma^* : 0w1 \in L\}.$$

Show that if L is regular, then so **is** H. To attain full credit, explicitly **construct** an FA M_H for H, given any FA M_L to recognize L; and formally **prove** that $L(M_H) = H$.

Question 9 (10 pts)

The following language is given.

$$L = \{1^t 0^m 1^{2^n} : t < 5,\ m \neq n, \text{ and } t, m, n \in \mathbb{N}\}.$$

Prove that L is **not** a regular language by

a. **pumping lemma** for regular languages. (10 pts)

b. MyHill-Nerode theorem. (not graded)

Question 10 (not graded)

Decide whether the following languages are **regular or not**. If you think a language is regular prove your claim by providing a regular expression or FA for it. Otherwise, use MyHill-Nerode theorem or pumping lemma for regular languages to prove that the language is not regular.

a. $L_1 = \{xy \in \{a,b\}^* : |x| = 2|y|$, and x contains the substring aa at indices 3 and 4$\}$

b. $L_2 = \{xy \in \{a,b\}^* : |x| = 2|y|$, and x ends with the substring a and y begins with the substring $a\}$

Question 11 (not graded)

Given DFA $M = (K, \Sigma, \delta, q_0, F)$, in which $K = \{q_0, q_1, q_2, q_3, q_4\}$, $\Sigma = \{0, 1\}$, $F = \{q_2, q_4\}$, and the state transition function σ is tabulated below,

q	σ	$\delta(q,\sigma)$
q_0	0	q_1
q_0	1	q_2
q_1	0	q_1
q_1	1	q_3
q_2	0	q_0
q_2	1	q_4
q_3	0	q_1
q_3	1	q_4
q_4	0	q_3
q_4	1	q_2

a. Apply **state minimization** algorithm to M to yield an equivalent DFA with minimum number of states.

b. Write down **equivalence classes** of Σ^* partitioned by the DFA with minimal states as regular expressions. For simplicity you may use shorthand notations such as Σ and L s.t. $L = L(M)$ within the regular expressions.

List of Acronyms

FA finite automaton

FA finite automata

DFA deterministic finite automaton

DFA deterministic finite automata

NFA nondeterministic finite automaton

NFA nondeterministic finite automata

GFA generalized finite automaton

Submission

- You should submit your THE1 as a pdf file with the identifier `'the1.pdf'` on odtuclass. You can use the provided .tex template with appropriate modifications.
- Soft-copies should be uploaded strictly by the deadline.
- **Late Submission:** You have two days in total for late submission with penalties of 20 pts and 50 pts reduction in your grade for the first and second day, respectively. No further submissions are accepted.
- Do not submit solutions for not-graded questions. Yet solving them is advisable in studying for the midterm.

Regulations

1. **Cheating:** This take-home exam has to be completed and submitted **individually**. Teaming up, sharing solutions anywhere other parties might access and using work belonging to others as part or in whole are considered cheating. **We have zero tolerance policy for cheating**. People involved in cheating will be punished in accordance with the university regulations.
2. **Newsgroup:** You must follow the newsgroup (cow.ceng.metu.edu.tr) for discussions and possible updates on a daily basis. You are advised to initiate discussions on COW so that most parties will benefit.