Middle East Technical University Department of Computer Engineering

CENG 280

Formal Languages and Abstract Machines

Spring 2016-2017

Take Home Exam 1

Due date: March 24, 2017, Fri, 23:55

Objectives

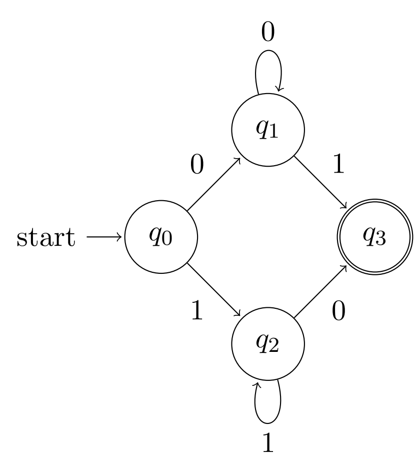
To familiarize with computation using the Finite Automata, the most restricted formal computational device, observing the benefits and limitations of employing such simplistic theoretical devices to tackle practical problems.

Specifications

You must adhere to the notation conventions adopted in the textbook. All finite 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 .tex file based on your modification of the provided template file. 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 (20 pts)

In each part, **draw** the state diagram of the DFA that accepts the given language and explicitly write down series of configurations one step at a time for strings w_t, stating whether w_t belongs to the language of the DFA or not. The alphabet is {0,1} in all parts.

a. $L_1 = \{$w | There is a 1 immediately after each 0 in w.$\}$ and $w_t = 11011011$.

b. $L_2 = \{$w | w starts with 10 and has length divisible by 3.$\}$ and $w_t = 1000111$.

Question 2 (10 pts)

Identify the language $L_3 = \{$w | w contains at least two 0's and ends with a 1.$\}$ as a regular expression.

Question 3 (10 pts)

Prove whether the following sets are countable or uncountable.

a. The set of configurations of a finite automaton on all strings over all possible finite alphabets

b. The set of configurations of all finite automata on all strings over all possible finite alphabets

Question 4 (20 pts)

Given NFA N=(K, Σ, Δ, q_0, F), in which $K = \{q_0, q_1, q_2, q_3, q_4\}$, $\Sigma = \{a, b\}$,
$\Delta = \{(q_0, a, q_0), (q_0, a, q_3), (q_0, b, q_1),$
$(q_1, a, q_2), (q_1, b, q_1), (q_1, b, q_2), (q_1, e, q_4),$
$(q_2, a, q_3), (q_2, b, q_0), (q_3, b, q_0), (q_3, e, q_4)\}$ and $F = \{q_4\}$,

a. Use the subset construction algorithm to find an equivalent DFA M s.t. $L(M) = L(N)$ and **draw** M.

b. Generate a regular expression for L^+ where $L = L(M)$.

c. Formally define and **draw** a DFA Q s.t. $L(Q) = L(M)^R$, where ^{R} denotes the reversal operation.

Question 5 (15 pts)

Prove that if L_1, L_2, and L_3 are regular languages over the alphabet $\Sigma = \{a, b\}$, then so is $L = (L_1L_3 - L_2) \cup L_4$ in which $L_4 = \{w \,|\, w = bx(ab)^*, x \in L_2\}$, and '$-$' denotes the set difference operation. Construct an NFA N to recognize L by using the NFA's for L_1, L_2, L_3 and L_4. Make sure that your formal proof is sound and complete.

Question 6 (10 pts)

Assume that L_1 and $L_1 \cup L_2 \cup L_3$ are regular languages. Is L_2L_3 necessarily a regular language? Prove by construction or disprove by counterexample.

Question 7 (15 pts)

Consider a hypothetical metro line comprising of possibly infinite number of stations and within every train there exists a digital circuit of lamps functioning as information panel, on which a consecutive sequence of red lights stands for the path that has been traveled from the initial station up to and including the current station and the following sequence of green lights represents the path to the final destination. Every station is marked by larger circular LED lamps and the path within two stations is just a serial connection of tinier lamps that could be lit either red or green.

Consequently, the path that has been traveled would cover a union of line segments connecting adjacent stations as lamps that are lit red and the distance of this path can be measured in terms of the number of larger lamps it contains. In other words, the cardinality of larger lamps on the red line segment must be equal to the number of stations between the current station and the initial station, both ends included. Similarly, the path to be traveled until arrival at the final destination would be a line segment lit in green which contains as many large lamps as the number of stations between the final station and the consecutive station of the current one, again including both ends.

Your goal is to devise an abstract language model for all possible panel configurations that a passenger can witness whenever they got on the train at any station of the metro line. For this purpose, the path traveled must be represented by a string of 1's whose length is equal to the number of larger lamps on the red-lit segment of the information panel, followed by another string of 1's whose length is equal to the number of larger lamps on the green-lit segment standing for the path to be covered next. Additionally, the length of the metro line in terms of the number of stations it connects must be appended as another string of 1's designating the corresponding integer value. Since all these strings are concatenation of unary 1's, 0's must be placed in between as separating characters so as to form the string representation of a valid panel configuration.

As an example, a person getting on the train at the second station of an eight-station long metro line would observe a red segment of lights containing two stations followed by a green segment of lights covering six stations. This valid configuration of the information panel is represented as the string of 1101111110111111111 in our case. Check that the lengths of the largest substrings of 1's are equal to 2, 6, and 8 respectively.

Figure 1: A configuration of LED indicators within a metro line in Ankara.

Another example can be found in what is actually used as information panel in Koru-Kizilay metro line which can be visually tracked through Fig. 1, despite the fact that it is not an image of the highest quality. In this case, the metro line has 12 stations, and anyone getting on the train at Beytepe station would immediately observe the configuration of the panel on the given image. It indicates that four stations have been already passed through and one-way journey on this train will end after visiting eight more stations.

a. Restricting the input alphabet as $\Sigma = \{0, 1\}$, give a mathematical definition of the set of all strings that constitute valid configurations of the information panel that anyone getting on the train at any station of the possibly infinitely long hypothetical metro line would observe. In other words, define the formal language of all such strings.

b. Prove or disprove whether the formal language you have identified is regular or not.

Submission

- **Late Submission:** You have 3 days in total for late submission of all homeworks. All homeworks will be graded as normal during this period. No further late submissions are accepted.
- You should submit your THE1 as a .tex file. Please use the template provided on COW with appropriate modifications.
- Soft-copies should be uploaded strictly by the deadline.

Regulations

1. **Cheating: We have zero tolerance policy for cheating**. People involved in cheating will be punished according to the university regulations.
2. **Newsgroup:** You must follow the newsgroup (news.ceng.metu.edu.tr) for discussions and possible updates on a daily basis.