Middle East Technical University Department of Computer Engineering

CENG 280

Formal Languages and Abstract Machines

Spring 2023-2024

Homework 1

Question 1

Give formal proofs to state whether the following sets are finite, countably infinite or uncountably infinite. State any mapping explicitly and give clear references to known theorems if used.

a) Intersection of the set of all strings on the alphabet $\Sigma = \{a, b, c\}$ which are starting with ab, ending with bc and the set of all strings on the same alphabet ($\Sigma = \{a, b, c\}$) which include the substring *"aabbcc"*.

b) The set of all regular languages on the binary alphabet $\Sigma = \{0, 1\}$.

c) The set of all languages on the binary alphabet $\Sigma = \{0, 1\}$.(This is, $2^{\{0,1\}^*}$)
Hint: Cantor's Diagonal Argument, regular expressions, closure properties may be helpful.

Question 2

- $L_{01} = \{\omega \in \{a, b\}^* |\ \omega$ includes an odd number of aba substrings$\}$
- $L_{02} = \{\omega \in \{a, b\}^* |$ neither baa nor ab is a substring of $\omega\}$
- $L_{03} = \{\omega \in \{a, b, c\}^* |$ every c is directly preceded by a and followed by $b\}$

For each of the languages given above;

a) Write a regular expression that generates the language.

b) Formally define and draw a DFA that recognizes the language.

Question 3

a) Using subset construction algorithm [1], construct an equivalent DFA for each of the NFAs given below:

- $N_1 = \{K_1, \Sigma_1, \Delta_1, s_1, F_1\}$ where
  $K_1 = \{q_0, q_1, q_2, q_3, q_4\}, \Sigma_1 = \{a, b\}, s_1 = q_0, F_1 = \{q_4\}$ and
  $\Delta_1 = \{$
  $(q_0, a, q_0), (q_0, a, q_1), (q_0, b, q_0),$
  $(q_1, b, q_2), (q_1, b, q_3),$
  $(q_2, a, q_0), (q_2, \epsilon, q_3),$
  $(q_3, a, q_4),$
  $(q_4, a, q_4), (q_4, b, q_4)$
  $\}$

- $N_2 = \{K_2, \Sigma_2, \Delta_2, s_2, F_2\}$ where
  $K_2 = \{q_0, q_1, q_2, q_3, q_4\}, \Sigma_2 = \{a, b\}, s_2 = q_0, F_2 = \{q_3\}$ and
  $\Delta_2 = \{$
  $(q_0, b, q_1), (q_0, b, q_3), (q_0, \epsilon, q_2),$
  $(q_1, a, q_1),$
  $(q_2, a, q_2), (q_2, b, q_2), (q_2, a, q_4),$
  $(q_3, a, q_1), (q_3, \epsilon, q_4),$
  $(q_4, a, q_2), (q_4, a, q_3), (q_4, b, q_4)$
  $\}$

b) Then, employing yields in one step relation[2] ($\vdash$) between configurations, trace the string $\omega_1 =$ *"aba"* on the given NFAs and on the equivalent DFAs you have constructed. For each of those four finite automata, decide whether $\omega_2 =$ *"babb"* is accepted by that automaton or not.

Question 4

Give a DFA that recognizes the language;

$$\{\omega \in \{a, b\}^* : \omega \text{ does } \textbf{NOT} \text{ have } aba \text{ as a substring}\}$$

a) Formally define the DFA as a quintuple $M = \{K, \Sigma, \delta, s, F\}$. Give each of K, Σ, δ, s, F.

b) For this DFA, trace the input *abbaabab* and write the computation. Does the DFA accept the input?

Question 5

Let us define a language similar to Turkish.

- In this language, a word consists of constants that can be denoted as **c**, or 8 different vowels **a**, **e**, **ı**, **i**, **o**, **ö**, **u**, **ü**.
- Every word consists of one or more syllables.

[1] Check Theorem 2.2.1 and Example 2.2.4 in your textbook.
[2] Check Example 2.1.1 in your textbook for DFA and Example 2.2.1 for NFA.

- A syllable starts with either a constant or a vowel.
- There can be one and only one vowel in a syllable.
- There can be at most two consecutive constants in a syllable.
- A syllable cannot start with more than one constant.
- If the first syllable of the word consists of one of **a**, **ı**, **o**, **u** vowels,then the other syllabuses must contain these vowels also.
- Similarly, if the first syllable of the word consists of one of **e**, **i**, **ö**, **ü** vowels,then the other syllabuses must contain these vowels also.
- **o** and **ö** vowels can only be in the first syllable of the word.

a) Give a NFA recognizes the language of strings which are valid syllables.

b) Write the regular expression for the set of the valid words according to given description.