Theory Assignment I

Automata Theory Monsoon 2024, IIIT Hyderabad

August 20, 2024

Total Marks: 35 points Due date: **22/08/24 11:59 pm**

General Instructions: All symbols have the usual meanings (example: $\mathbb{R}$ is the set of reals, $\mathbb{N}$ the set of natural numbers, and so on). FSM stands for finite state machine. DFA stands for deterministic finite automata. NFA stands for non-deterministic finite automata. a^* is the Kleene Star operation.

---

1. [2 points] Given a NFA with 5 states, what is the maximum number of transitions it can have if the alphabet has 3 symbols? **[CO 1]**

2. [3 points] A Finite State Transducer (FST) is a 5-tuple $M = (Q, \Sigma, \Gamma, \delta, s)$ where-

   - Q is a finite set of states,
   - Σ is a finite set of input symbols
   - Γ is a finite set of output symbols
   - $\delta : Q \times (\Sigma \cup \{\epsilon\}) \longrightarrow Q \times (\Gamma \cup \{\epsilon\})$ is the transition function, allowing for epsilon transitions in both input and output,
   - $s \in Q$ is the start state.

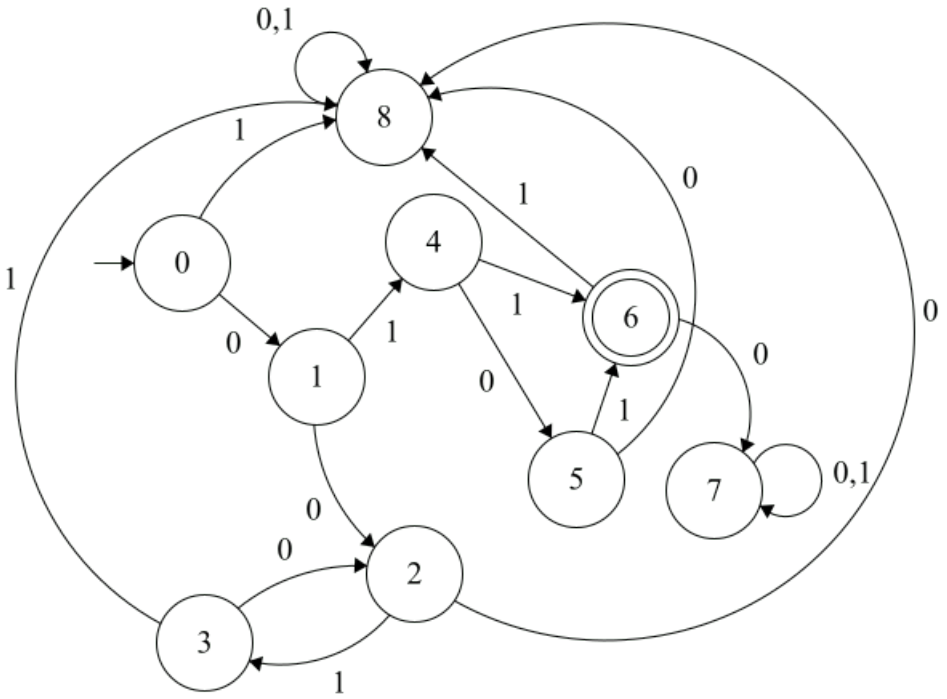
   Construct a transducer, that takes as input a program, and writes in the output the part of the program that is not commented. Σ = set of unicode characters

   - Every comment starts and ends with %%
   - If the input contains the start of a comment but not its end, then the entire program after the start of the comment is commented

   Eg :
   Input
   print(”hi”) %%testing%%
   print(123)
   %%this is a comment%% print(”this is not”) %%this is a comment again%%
   Output
   print(”hi”)
   print(123)
   print(”this is not”)

   [CO 1, CO 2]

3. [3 points] Minimize the following DFA.

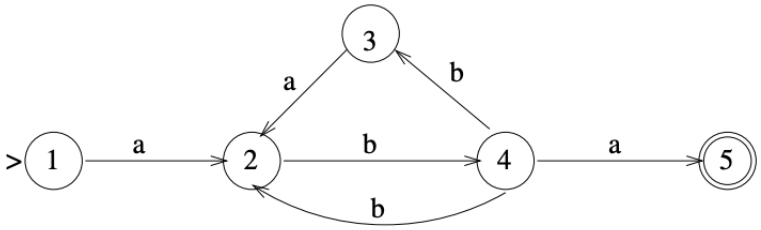


[CO 1, CO 2]

4. [3 points] Design a NFA to accept the set of strings of 0's and 1's that either
(a) end in 010 and have 011 somewhere preceding, or
(b) end in 101 and have 100 somewhere preceding

[CO 1, CO 2]

5. [2 points] (i) What is the language accepted by the following finite state automata?



(ii) Let $\Sigma = \{a, b\}$. Write regular expression for the language L consisting of all strings in Σ^* with exactly one occurrence of the substring aaa. **[CO 1, CO 2]**

6. [3 points] If A is any language, let $A_{1/3-1/3}$ be the set of all strings in A with their middle thirds removed so that $A_{1/3-1/3} = \{xz \mid \text{for some } y, |x| = |y| = |z| \text{ and } xyz \in A\}$. Show that if A is regular, then $A_{1/3-1/3}$ is not necessarily regular. **[CO 1, CO 2, CO 3, CO 4]**

7. [3 points] Let M_1 and M_2 be DFAs that have k_1 and k_2 states, respectively, and then let $U = L(M_1) \cup L(M_2)$. Show that if $U \neq \emptyset$, then U contains some string s, where $|s| < \max(k_1, k_2)$. **[CO 1, CO 2]**

8. [3 points] Show that the following languages are not regular using the pumping lemma:

   1. $L = \{w \mid w \text{ has balanced parentheses}\}$
   2. $L = \{a^{n!} \mid n \geq 0\}$ **[CO 1, CO 2, CO 3, CO 4]**

9. [2 points] Let the grammar G of L be the one below:

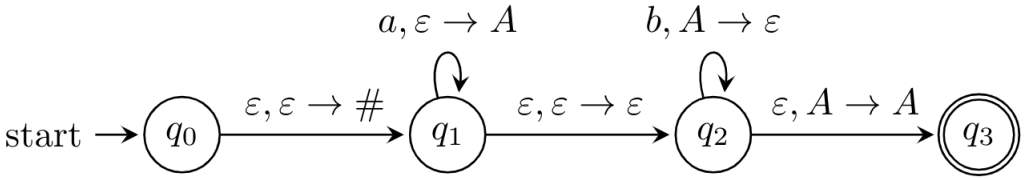
$$R \to ST \mid UV$$
$$T \to UV \mid W$$
$$V \to XY \mid Z$$
$$X \to YZ \mid T$$

   The above grammar G is one whose variables and terminals are NOT named using the usual convention. Any of the symbols R through Z could be either a variable or a terminal; your task is to point out which is which, and which could be the start symbol. **[CO 1, CO 2]**

10. [3 points] Give a CFG to generate $A = \{a^i b^j c^k \mid i, j, k \geq 0 \text{ and either } i = j \text{ or } j = k\}$. Is the grammar ambiguous? Why or why not? **[CO 1, CO 2, CO 3, CO 4]**

11. [4 points] Construct a PDA for the language of all non-palindromes over {a, b}. **[CO 1, CO 2]**

12. [4 points] Consider the following PDA over the input alphabet $\Sigma = \{a, b\}$ and stack alphabet $\Gamma = \{\#, A\}$



   Describe the language decided by the given PDA and then find the number strings of length 100 accepted by it. **[CO 1, CO 2]**