Quiz 1 Solutions

Automata Theory Monsoon 2021, IIIT Hyderabad

November 18, 2021

Total Points: 20

General Instructions: FSM stands for finite state machine. DFA stands for deterministic finite automata. NFA stands for non-deterministic finite automata. PDA stands for Push Down Automata. a^* is the Kleene Star operation. $a^+ = a^* \setminus \{\epsilon\}$, where ϵ is the empty string.

---

1. [2 points] Proof:

   - If L is a language containing a finite number of strings $a_0, a_1, \cdots, a_n$.
   - The language $\{a_i\}$ consisting of a single literal string a_i is regular.
   - The union of a finite number of regular languages is also regular.
   - Therefore, $L = \{a_0\} \cup \{a_1\} \cup \cdots \cup \{a_n\}$ is regular.

   Note: Marks have also been given for writing the final regular expression given above.

   OR

   Using the induction method, where first it is proven that the language containing a single string is regular, the language containing n strings is assumed to be regular, and then the union of these two languages is proven to be regular.

   OR

   Constructing a DFA/NFA with the correct notations being followed.

2. [4 points] Construct a NFA $M = \{Q, \Sigma, \delta, q, F\}$ using the NFAs for B and C, namely,

   $$M_B = \{Q_B, \Sigma, \delta_B, q_B, F_B\}$$

   $$M_C = \{Q_C, \Sigma, \delta_C, q_C, F_C\}$$

   can be defined as:

   - $Q = Q_B \times Q_C$
   - $F = F_B \times F_C$
   - $q = (q_B, q_C)$, with

   $$(\delta(q,r), a) = \begin{cases} (\delta_B(q,0), r) & \text{if } a = 0 \\ (\delta_B(q,1), \delta_C(r,1)) & \text{if } a = 1 \\ (q, \delta_C(r,0)) & \text{if } a = \epsilon \end{cases}$$

   where $(q,r) \in Q$ and $a \in \Sigma$ (with suitable explanation supporting the above construction)

- 4 for defining each of Q, F, q and δ correctly
- 3.5 for missing/erroneously defining either of Q, F or q (0.5 deducted for missing each of these three)
- 1 for stating a procedure without formally defining M in terms of Q, F, q and δ
- alternative solutions are graded according to their merit

3. [2 points] Chomsky Normal form of the given language is:

$$S \to AB$$

$$A \to a|BU$$

$$B \to BU$$

$$U \to b$$

There are total 5 production rules.

- 1 for removing the rule: $A \to B$.
- 0.5 for removing the rules: $A \to Bb$ and $B \to Bb$.
- 0.5 for writing the number of productions.

4. [2 points] Consider the string $w = aacbc$. There are two different parse trees for this string. The following are the two leftmost derivations:

$S \to aSbS \to aaSbS \to aacbS \to aacbc,$ [1]

$S \to aS \to aaSbS \to aacbS \to aacbc.$ [1]

Hence, G is ambiguous.
Note: Alternative strings with correct parse trees have also been given marks. In case logical steps haven't been provided to the answer, 0.25 - 0.5 marks have been deducted.

5. [3 points] L is regular and is simply strings that begin and end with the same alphabet. The corresponding regular expression is:

$$a(a+b)^+a + b(a+b)^+b.$$

- Zero marks for writing non regular
- One mark for writing regular
- Two marks for drawing the DFA or writing the regex.

6. [3 points] The grammar G generating L has the following rules:

$$S \to 0A1111$$
$$A \to 0A1111|\epsilon$$

The PDA would have the following transition functions:

$$\begin{aligned} \delta(q_0, 0, \epsilon) &= (q_0, XXXX) \\ \delta(q_0, 1, X) &= (q_1, \epsilon) \\ \delta(q_1, 1, X) &= (q_1, \epsilon) \\ \delta(q_1, \epsilon, \$) &= (q_2, \$). \end{aligned}$$

0.25 for recognizing that it's context free, 0.75 for correct grammar, 2 for correct PDA

7. [4 points] (i) L_1 is not regular. It is context-free. The PDA for L_1 is given below:

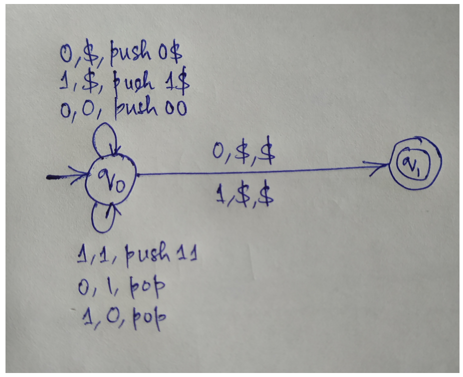

Figure 1: PDA for L_1

(ii) L_2 is regular. The DFA for L_2 is given below

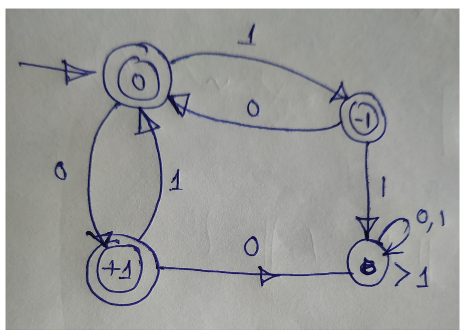

Figure 2: DFA for L_2

Comments from correction:

- Marks distribution: 1 mark for correctly mentioning Context free or Regular with some justification. 1 mark for the automaton. In part 2, mentioning CF has been given 0.5, 1 if you mention Regular.
- A good percentage of people have tried to argue that since $L = \{w|w = 0^n 1^n : n \in \mathbb{N}\} \subset L_1$, it is context free. This is incorrect. Note that this is incorrect, since I can use a similar argument to prove Σ^* is context free.

- A good percentage of students seem to have copied one answer where they try to incorrectly build a CFG for the languages, which has been given a 0. They later don't show if the languages are Regular or not, Context free or not.
- A lot of people have assumed L_2 is just alternating 0s and 1. Note that $01011 \in L_2$. Similarly, a lot of people have incorrectly assumed L_1 only has strings of the form n^n1^n or 1^n0^n.
- Just blindly guessing Context free or regular for both is given 0 marks. If a student has given some indication that they have thought something close to the correct lines, they have been given marks for mentioning. If they go on to try and construct PDA or DFA, they have been given full marks for the mentioning part, even if the automaton was incorrect. Marks have been deducted for incorrect automaton appropriately.