1 Short Answer (Regular) (2 parts) - 20 points

No explanation is required for your answers for full credit. Keep any explanations of your answers to 2 sentences maximum.

a. Give the regular expression for the following language:

- All strings that contain **010** or **101** as a suffix.

Solution: If w is a string and w belongs to the language L, we can split w into two parts $w = xy$ such that y represents the suffix which can be **010** or **101** and x is the prefix which can be any string. We can write $x = (0 + 1)^*$ and y as $(010 + 101)$ Therefore the regular expression for the language is

$$(0 + 1)^* \, (010 + 101)$$

■

- The language that **does not** contain **10** as a sub-sequence.

Solution: To generate a language that does not have **10** as a sub-sequence, there can be no **0** appearing after the first occurrence of **1**. This means that all occurrences of **0**s should be before a **1**. So the string can be ε, **0**, **1**, **00**, **01**, **11**, **000**, **001**, **011**, **111** and so on. We can see that all **0**s occur before the first **1** or the string consists only of **0**s. The regular expression can be given as:

$$0^* \, 1^*$$

■

b. Construct the DFA that describes the following language:

$$L = \{w \in \Sigma^* | w \text{ has a even number of } 1\text{'s } \textbf{and} \text{ the substring } 00\}$$

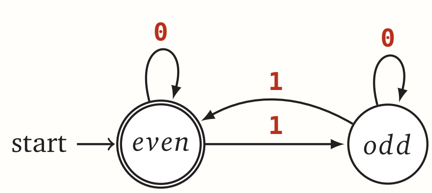
(you can draw it out, or describe it formally)

Solution: Let us consider the two languages:

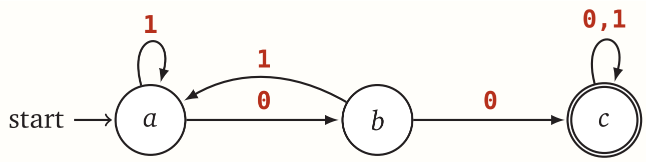
$$L_1 = \{w_1 \in \Sigma^* | w_1 \text{ has a even number of } 1\text{'s} \}$$

$$L_2 = \{w_2 \in \Sigma^* | w_2 \text{ has the substring } 00\}$$

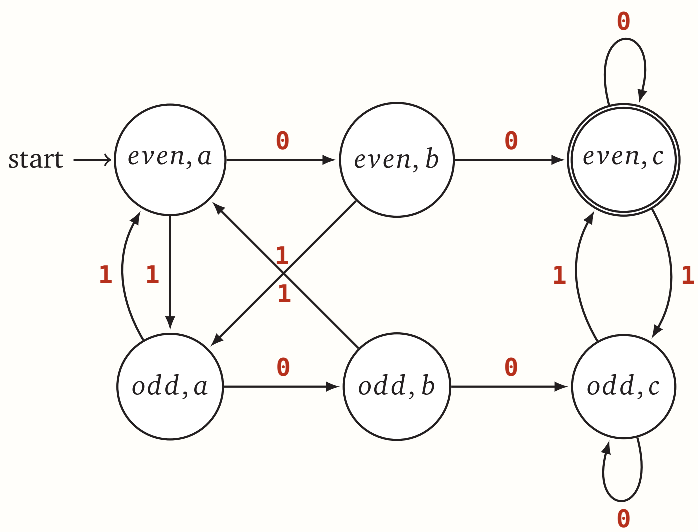
DFA for L_1:



DFA for L_2:



Constructing the DFA by the Cross Product of these two DFAs:



Let

$$M_1 = (Q_1, \Sigma, \delta_1, s_1, A_1) \text{represent the DFA that accepts} L_1$$

and

$$M_2 = (Q_2, \Sigma, \delta_2, s_2, A_2) \text{represent the DFA that accepts} L_2$$

Formal definition can be given by

$$M = (Q, \Sigma, \delta, s, A)$$

where:

$$Q = Q_1 \times Q_2$$
$$s = (s_1, s_2)$$
$$A = (q_1, q_2) \text{ where } q_1 \epsilon A_1 \text{ and } q_2 \epsilon A_2$$
$$\delta((q_1, q_2), a) = (\delta_1(q_1, a), \delta_2(q_2, a))$$

■

2 Short Answer (Context-Free) (2 parts) - 20 points

No explanation is required for your answers for full credit. Keep any explanations of your answers to 2 sentences maximum.

a. Consider the inductive definition of a language MYSTERY:

- $\texttt{0} \in$ MYSTERY
- If $w \in$ MYSTERY, then $\texttt{1}w\texttt{1} \in$ MYSTERY
- If $w \in$ MYSTERY, then $\texttt{0}w\texttt{0} \in$ MYSTERY

Give the context-free-grammar for this language

Solution: $S \rightarrow \texttt{0} \mid \texttt{0}S\texttt{0} \mid \texttt{1}S\texttt{1}$

The center symbol must be a 0 then we can add 0s or 1s on either side. ■

b. Consider the following context free grammar:

$$S \rightarrow AB \mid B$$
$$A \rightarrow \epsilon \mid \texttt{0}A$$
$$B \rightarrow \texttt{1}B\texttt{2} \mid \texttt{12}$$

Describe the language that the above CFG represents (Ex. $L = \{\ldots\}$)

Solution: $L = \{\texttt{0}^m\texttt{1}^n\texttt{2}^n \mid m \geq 0, n > 0\}$

A generates an arbitrary amount of 0s. *B* generates an an arbitrary amount of 1s (at least one) followed by an equal amount of 2's. ■

3 Language Transformation - 20 points

Assume L is a regular language.

Prove that the language $DeleteA(L) := \{xy \mid x0y \in L \text{ or } x1y \in L\}$ is regular.

Intuitively $DeleteA(L)$ is every string which can be made by deleting a character from a string in L. So if $L = \{\varepsilon, \texttt{010}\}$, then $DeleteA(L) = \{\texttt{10}, \texttt{00}, \texttt{01}\}$

Solution: Let $M = (\Sigma, Q, \delta, s, A)$ be a DFA for the language L. We can construct an NFA $M' = (\Sigma, Q', \delta', s', A')$ for $DeleteA(L)$ as the following:

- $Q' = Q \times \{pre, post\}$
- $s' = (s, pre)$
- $A' = \{(f, post) \mid f \in A\}$
- $\delta' : Q' \times \Sigma \to Q'$ is defined as the following.

$$\delta'((q, pre), a) = \{(\delta(q, a), pre)\}, \qquad a \neq \epsilon$$
$$\delta'((q, post), a) = \{(\delta(q, a), post)\}, \qquad a \neq \epsilon$$
$$\delta'((q, pre), \epsilon) = \{(\delta(q, a), post) \mid a \in \Sigma\}$$

The deletion of a symbol must occur exactly once, so we have another copy of the original states to indicate that the deletion has occurred. The deletion is done by taking an epsilon transition to the next state in the post-deletion copy. Since we could construct an NFA for $DeleteA(L)$, we conclude that $DeleteA(L)$ is regular. ■

4 Language classification I (2 parts) - 20 points

Let $\Sigma = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ and

$$L_4 = \{w \in \Sigma^* | \text{ the bottom row of } w \text{ is the reverse of the top row of } w\}.$$

For instance: $\begin{bmatrix} 1 \\ 0 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \in L_4$ but $\begin{bmatrix} 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \notin L_4$

1. Is L_4 regular? Indicate whether or not by circling one of the choices below. Either way, prove it.

Solution: regular (not regular)

The given language L_4 is irregular and so let's try to prove it through fooling sets.

Let F $= \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}^n \mid n \geq 0 \right\}$. We take two arbitrary strings x and y from our fooling set such that x $= \begin{bmatrix} 1 \\ 0 \end{bmatrix}^i$ and y $= \begin{bmatrix} 1 \\ 0 \end{bmatrix}^j$ where i and j are two non-negative integers. If suffix z $= \begin{bmatrix} 0 \\ 1 \end{bmatrix}^i$, then xz $= \begin{bmatrix} 1 \\ 0 \end{bmatrix}^i \begin{bmatrix} 0 \\ 1 \end{bmatrix}^i$ and yz $= \begin{bmatrix} 1 \\ 0 \end{bmatrix}^j \begin{bmatrix} 0 \\ 1 \end{bmatrix}^i$. In xz, for every 0 in the bottom row of x there is a 0 on the top row of z and for every 1 on the top row there is a 1 on the bottom row in z. Since we have equal pairs of corresponding the top row of xz is essentially the reverse of the bottom row, xz $\in L_4$. But when it comes to yz, since j $\neq$ i, there are no equal pairs of 0s and 1s in y and z and hence the bottom row is not the reverse of the top row proving that yz $\notin L_4$. Hence we have proved that F is a valid fooling set for L_4 and since F is an infinite fooling set, the language L_4 is irregular. ■

2. Is L_4 context-free? Indicate whether or not by circling one of the choices below. Either way, prove it.

Solution: (context-free) not context-free

The given language is context-free and so lets construct a context-free grammar for the same. If we inspect any string s from L_4, every pair of the left outermost and the right outermost symbols, is essentially a 'reverse' of one another and this true for all pairs, from the outermost to the innermost. That is, each pair consists of either $\begin{bmatrix} 1 \\ 0 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, all 0s or all 1s. Additionally, in case of strings with odd length the symbol left at the center can be $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ or $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.So the context-free grammar would be :

$$S \rightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix} S \begin{bmatrix} 0 \\ 1 \end{bmatrix} \Big| \begin{bmatrix} 0 \\ 1 \end{bmatrix} S \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Big| \begin{bmatrix} 1 \\ 1 \end{bmatrix} S \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Big| \begin{bmatrix} 0 \\ 0 \end{bmatrix} S \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Big| \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Big| \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Big| \epsilon$$

■

5 Language classification II (2 parts) - 20 points

Let $\Sigma = \{0, 1\}$ and

$$L_5 = \{0^n w 1^n \mid w \in \Sigma^*, n \geq 0\}.$$

1. Is L_5 regular? Indicate whether or not by circling one of the choices below. Either way, prove it.

Solution: (regular) not regular

The language is regular. When $n = 0$ the whole string is just represented by w. In this case all the strings over $\{0, 1\}$ can be just represented by w as $w \in \Sigma^*$. So,we can ignore the 0^n,1^n portion as every string is covered by w. Therefore, every string in L_5 can be represented by w.

The regular expression for L_5 will be $(0 + 1)^*$.

Hence,the language L_5 is a regular language. ■

2. Is L_5 context-free? Indicate whether or not by circling one of the choices below. Either way, prove it.

Solution: (context-free) not context-free

The language is context-free as all regular languages are context free. Since, w can define every string in the language, we just need to define the context-free grammar which would cover every string over Σ^*.

$S \rightarrow \varepsilon \mid 0S \mid 1S$

Hence,the language L_5 is context free. ■