1. Short answer type question. Please give appropriate reason wherever necessary.

   (a) **(1 mark)** What is language recognised by the regular expression $\emptyset^*$?

   (b) **(2 marks)** Every language recognised by a regular expression, is also generated by some CFG. True or false.

   (c) **(4 marks)** If L_1 and L_2 are two CFLs, then is $\overline{L_1}$ a CFL? What about $L_1 \setminus L_2$?

   (d) **(3 marks)** Consider an NFA $N = (Q, \Sigma, \delta, q_0, Q)$ such that $|\delta(q, a)| > 0$ for all $q \in Q$ and $a \in \Sigma$. Is $L(N) = \Sigma^*$?

Solution:

(a) $\{\epsilon\}$.

(b) True since regular languages are a subclass of CFLs.

(c) Both are false. We discussed in class that CFLs are closed under union and there exists CFLs L_1 and L_2, such that $L_1 \cap L_2$ is not a CFL.

Now we can write,

$$\begin{aligned} L_1 \cap L_2 &= \overline{\overline{L_1} \cup \overline{L_2}} \quad \text{and} \\ L_1 \cap L_2 &= L_1 \setminus (L_2 \setminus L_1). \end{aligned}$$

Therefore if CFLs are closed under complement or set difference, then they are closed under intersection as well. This is a contradiction.

(d) Yes.

Observe that (i) every state in N is an accept state, and, (ii) for every state $q \in Q$ and for every symbol $a \in \Sigma$, there is at least one transition from q on a. Therefore, for every string $w \in \Sigma^*$, there is at least one sequence of states in N that corresponds to the sequence of symbols in w. Therefore $w \in L(N)$.

□

2. **(5 marks)** Give a regular expression for the language

$$L = \{w \in \{0,1\}^* \mid w \text{ has an odd number of 1's}\}.$$

Solution:

$$0^*(10^*10^*)^*10^*$$

□

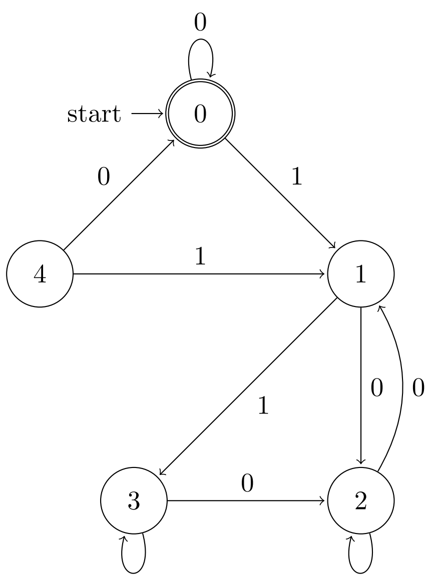
3. Let $Q = \{0, 1, 2, 3, 4\}$. Consider the DFA $D = (Q, \{0, 1\}, \delta, 0, \{0\})$, where $\delta(q, i) = (q^2 + q + i) \bmod 5$.

   (a) **(4 marks)** Construct the state diagram of D.

   (b) **(5 marks)** Minimise D.

   (c) **(2 marks)** Give the simplest possible description of the set of strings recognised by D.

Solution:



(a)

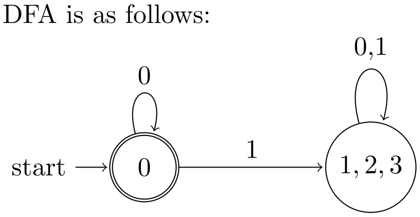
(b) Initial table (Note that we do not include 4 since 4 is not reachable from the start state)

0			
-	1		
-	-	2	
-	-	-	3

After first pass (marked all pairs consisting of one accept state and one non-accept state).

0			
×	1		
×	-	2	
×	-	-	3

At this stage no further new pairs can be marked, hence the algorithm terminates. The minimised DFA is as follows:



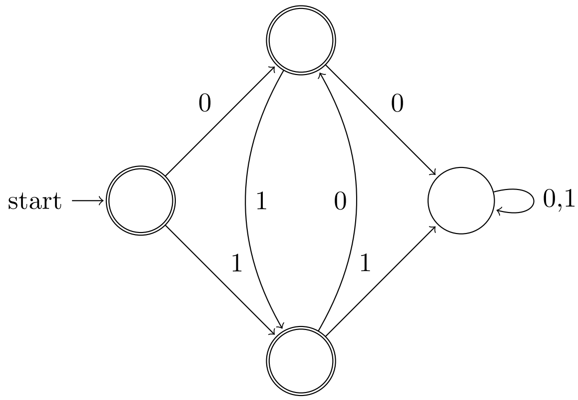
(c)

$$L(D) = \{w \in \{0, 1\}^* \mid w \text{ does not contain any } 1\}.$$

□

4. **(8 marks)** Construct a DFA for the language

$$L = \{w \in \{0,1\}^* \mid w \text{ neither contains } 00 \text{ nor } 11 \text{ as a substring}\}.$$



Solution: □

5. Consider the language

$$L = \{0^i 1^j \mid 0 \leq 2i \leq j\}.$$

(a) **(5 marks)** Prove that L is not regular.

(b) **(5 marks)** Give a CFG for L.

(c) **(4 marks)** Give a parse tree for the string $w = 0^2 1^7$ with respect to the grammar you constructed in part (b).

Solution:

(a) We prove L is not regular using the contrapositive form of the pumping lemma for regular languages. Given $p > 0$, pick $w = 0^p 1^{2p}$. Clearly $w \in L$. Now for any partition $w = xyz$, consider the string $w' = xy^2z$. Since $|xy| \leq p$, y consists only of 0's. Let $y = 0^t$. Therefore w' is of the form

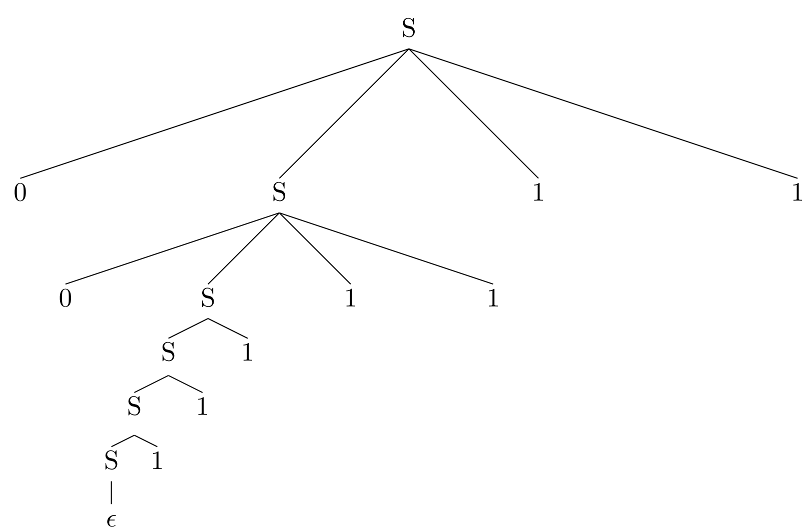
$$w' = 0^{p+t} 1^{2p}.$$

Now since $t > 0$ therefore $2(p+t) > 2p$, and thus, $w' \notin L$.
Hence L is not regular.

(b)

$$S \longrightarrow 0S11 \mid S1 \mid \epsilon$$

(c)



□

6. **(7 marks)** Is the following language context-free? Prove or disprove.

$$L = \{a^i b^j c^k \mid i \geq j, i \geq k, i, j, k \geq 0\}.$$

Solution: We prove L is not context-free using the contrapositive form of the pumping lemma for CFLs. Given $p > 0$, pick $w = a^p b^p c^p$. Clearly $w \in L$. Now for any partition $w = uvxyz$, such that $|vxy| \leq p$ and $|vy| > 0$, consider the following cases:

- (Case 1: vy has some a's) Since vy has some a's, vy cannot have any c's. Therefore consider the string $w' = uv^0xy^0z$. w' has strictly lesser number of a's than c's. Hence $w' \notin L$.
- (Case 2: vy has no a's) Since $|vy| > 0$ and it contains no a's, therefore it must contain some b's or c's. Consider the string $w' = uv^2xy^2z$. w' has strictly greater number of b's or c's than a's. Hence $w' \notin L$.

Hence L is not context-free.

□

7. **(10 marks)** Give a CFG for the language

$$L = \{a^i b^j c^k \mid j \le i + k \le 2j, i, j, k \ge 0\}.$$

Solution:

$$\begin{aligned}
S &\longrightarrow S_1 S_2 \mid aS_1 bS_2 c \\
S_1 &\longrightarrow AS_1 b \mid \epsilon \\
A &\longrightarrow a \mid aa \\
S_2 &\longrightarrow bS_2 C \mid \epsilon \\
C &\longrightarrow c \mid cc
\end{aligned}$$

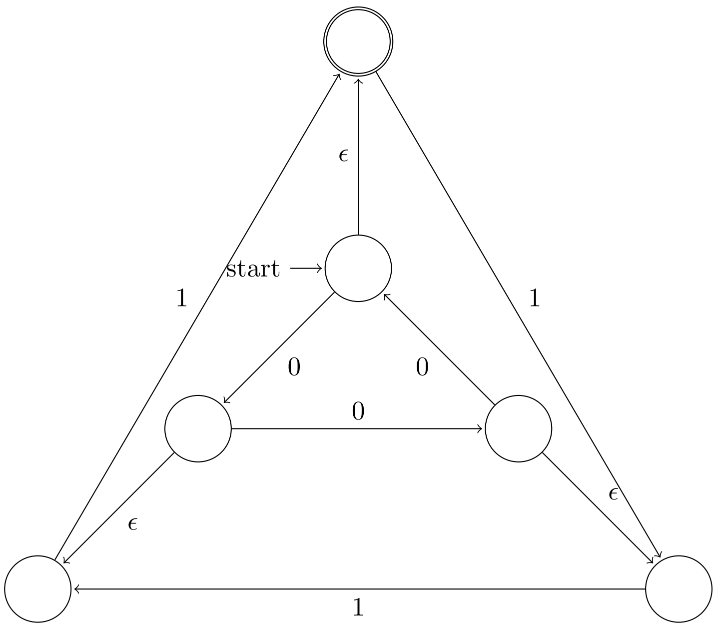
□

8. **(10 marks)** Is the following language regular? Prove or disprove.

$$L = \{0^{n+3k}1^n \mid n \geq 0, k \in \mathbb{Z}\}.$$

Solution: Observe that, we can rephrase L as set of strings of the form $0^i 1^j$ such that i and j leave the same remainder modulo 3.

Here is an NFA for L.



□

9. **(15 marks)** For a language $L \subseteq \{0,1\}^*$, define

$$\text{cycle}(L) = \{yx \mid xy \in L \text{ and } x, y \in \{0,1\}^*\}.$$

For example, if $L = \{00, 01, 101\}$ then $\text{cycle}(L) = \{00, 01, 10, 101, 011, 110\}$.

Show that if L is regular then cycle(L) is regular. If your answer involves a construction, give appropriate justification for your construction.

Solution:

Idea

We will construct an NFA N that would recognise cycle(L). The idea of N is to simulate a pebble game with three pebbles (say R, G and B), such that for a string $w = xy$ in L, B will simulate the string y, R will simulate the string x and G is used as a marker to remember the starting position of B. We will also have a fourth component (a binary value) to record the fact whether or not we have finished simulating x (and therefore should move onto x).

The NFA N that we will construct will have several start states, with the proper that a string is accepted by N if it reaches an accept state while starting at "some" start state. This assumption can be made without loss of generality as discussed in class.

Formal Construction

Let $D = (Q, \{0,1\}, \delta, q_0, F)$ be a DFA that such that $L = L(D)$.

Define

$$\begin{aligned} Q' &:= Q^3 \times \{0,1\} \\ S' &:= \{(q_0, q, q, 0) \mid q \in Q\} \\ F' &:= \{(q, q, f, 1) \mid q \in Q, f \in F\} \\ \delta'((p,q,r,b),a) &\ni \begin{cases} (p, q, \delta(r,a), b) & \text{if } b = 0 \\ (\delta(p,a), q, r, b) & \text{if } b = 1 \\ (p, q, r, 1) & \text{if } r \in F \text{ and } a = \epsilon \text{ and } b = 0 \end{cases} \end{aligned}$$

Define the NFA N as $N = (Q', \{0,1\}, \delta', S', F')$.

Claim 1. *cycle*$(L) = L(N)$

Proof. ($\Longrightarrow$) Let $w \in \text{cycle}(L)$. Then there exists strings x, y such that $w = xy$ and $yx \in L$. Let $p = \delta(q_0, y)$. Then observe that $\delta(p, x) f(\text{say}) \in F$.

Now consider the state $(q_0, p, p, 0) \in S'$. By our earlier observation we have $\delta'((q_0, p, p, 0), x) \ni (q_0, p, \delta(p, x), 0) = (q_0, p, f, 0)$. Since $f \in F$, we also have $\delta'((q_0, p, f, 0), \epsilon) \ni (q_0, p, f, 1)$. Finally, $\delta'((q_0, p, f, 1), 1) \ni (\delta(q_0, y), p, f, 1) = (p, p, f, 1) \in F'$. Therefore, combining the three we get, $\delta'((q_0, p, p, 0), xy) \ni (p, p, f, 1) \in F'$. Hence, $xy = w \in L(N)$.

($\Longleftarrow$) Let $w \in L(N)$. Therefore there exists states $(q_0, q, q, 0) \in S'$ and $(q, q, f', 1) \in F'$ such that $(q, q, f', 1) \in \delta'((q_0, q, q, 0), w)$.

Since the last component of the two states are different there must exist a state of the form $(q_0, q, f', 0)$ (for some $f' \in F$), where this flip occurs. Moreover, by the definition of δ' the bit can only be flipped from 0 to 1, and therefore such a flip occurs exactly once along a computation path of N. Also note that the state at which the flip occurs, must have its first component as q_0 as the first component does not change when the fourth component is 0, and it must have its third component as f', because once the flip occurs, the third component never changes according to the definition of δ'.

Let x_1 be the prefix of w such that $(q_0, q, f', 0) \in \delta'((q_0, q, q, 0), x_1)$, and let $w = x_1 x_2$. Then $(q, q, f', 1) \in \delta'((q_0, q, q, 0), x_2)$. This means that $\delta(q, x_1) = f'$ and $\delta(q_0, x_2) = q$. Combining the two we have $\delta(q_0, x_2 x_1) = f' \in F$, which implies that $x_2 x_1 \in L$ and therefore $x_1 x_2 = w \in \text{cycle}(L)$. □

□