Suggested Solutions to Midterm Problems

1. (a) Give a binary relation on $A = \{a, b, c, d\}$ that is symmetric and transitive but not reflexive.

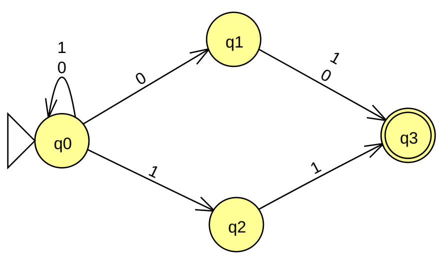
Solution. $R = \{(a, c), (c, a), (a, a), (c, c)\}$ is a binary relation on A that is symmetric and transitive but not reflexive (for example, $b \in A$ but $(b, b) \notin R$). □

(b) Let $R = \{(a, c), (b, c), (d, e)\}$ be a binary relation on $A = \{a, b, c, d, e\}$. Find the smallest equivalence relation on A that contains R.

Solution. $R' = \{(\mathbf{a}, \mathbf{a}), (\mathbf{b}, \mathbf{b}), (\mathbf{c}, \mathbf{c}), (\mathbf{d}, \mathbf{d}), (\mathbf{e}, \mathbf{e}), (a, c), (\mathbf{c}, \mathbf{a}), (b, c), (\mathbf{c}, \mathbf{b}), (\mathbf{a}, \mathbf{b}), (\mathbf{b}, \mathbf{a}), (d, e), (\mathbf{e}, \mathbf{d})\}$ is the smallest equivalence relation on A that contains R. □

2. (a) Draw the state diagram of an NFA, with as few states as possible, that recognizes $\{w \in \{0, 1\}^* \mid w \text{ ends with } 00, 01, \text{ or } 11\}$. The fewer states your NFA has, the more points you will be credited for this problem. (5 points)

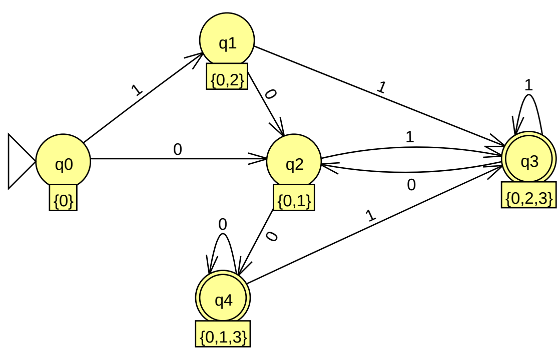
Solution.



□

(b) Convert the NFA in (a) systematically into an equivalent DFA (using the procedure discussed in class); do not attempt to optimize the number of states. (10 points)

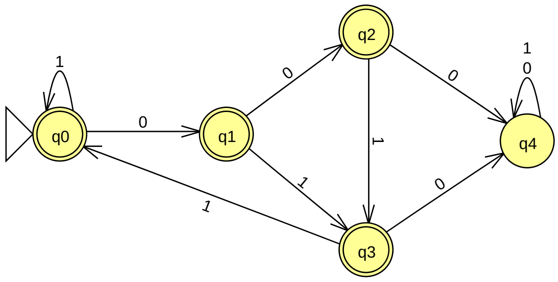
Solution.



Unreachable states have been removed. □

3. (a) Draw the state diagram of a DFA, with as few states as possible, that recognizes $\{w \in \{0,1\}^* \mid w \text{ doesn't contain 000 or 010 as a substring}\}$. The fewer states your DFA has, the more points you will be credited for this problem. (10 points)

Solution.



□

(b) Translate the DFA in (a) systematically to an equivalent context-free grammar (using the procedure discussed in class). (5 points)

Solution.

$$
\begin{array}{lcl}
Q_0 & \to & 0Q_1 \mid 1Q_0 \mid \varepsilon \\
Q_1 & \to & 0Q_2 \mid 1Q_3 \mid \varepsilon \\
Q_2 & \to & 0Q_4 \mid 1Q_3 \mid \varepsilon \\
Q_3 & \to & 0Q_4 \mid 1Q_0 \mid \varepsilon \\
Q_4 & \to & 0Q_4 \mid 1Q_4
\end{array}
$$

□

4. Write a regular expression for the language in Problem 3.

Solution. To be provided; you may use JFLAP to quickly find an answer. □

5. Convert the following CFG into an equivalent CFG in Chomsky normal form (using the procedure discussed in class: add a new start symbol, remove illegal ε rules, remove unit rules, ...).

$$
\begin{array}{lcl}
A & \to & BAB \mid B \mid 11 \\
B & \to & 00 \mid \varepsilon
\end{array}
$$

Solution.

(a) Add a new start symbol S.

$$
\begin{array}{lcl}
S & \to & A \\
A & \to & BAB \mid B \mid 11 \\
B & \to & 00 \mid \varepsilon
\end{array}
$$

(b) Remove ε-rule $B \to \varepsilon$.

$$
\begin{array}{lcl}
S & \to & A \\
A & \to & BAB \mid B \mid 11 \mid AB \mid BA \mid A \mid \varepsilon \\
B & \to & 00
\end{array}
$$

(c) Remove ε-rule $A \to \varepsilon$.

$$\begin{array}{lcl} S & \to & A \mid \varepsilon \\ A & \to & BAB \mid B \mid 11 \mid AB \mid BA \mid A \mid BB \\ B & \to & 00 \end{array}$$

(d) Remove unit rule $S \to A$.

$$\begin{array}{lcl} S & \to & BAB \mid B \mid 11 \mid AB \mid BA \mid BB \mid \varepsilon \\ A & \to & BAB \mid B \mid 11 \mid AB \mid BA \mid A \mid BB \\ B & \to & 00 \end{array}$$

(e) Remove unit rule $S \to B$.

$$\begin{array}{lcl} S & \to & BAB \mid 00 \mid 11 \mid AB \mid BA \mid BB \mid \varepsilon \\ A & \to & BAB \mid B \mid 11 \mid AB \mid BA \mid A \mid BB \\ B & \to & 00 \end{array}$$

(f) Remove unit rule $A \to A$.

$$\begin{array}{lcl} S & \to & BAB \mid 00 \mid 11 \mid AB \mid BA \mid BB \mid \varepsilon \\ A & \to & BAB \mid B \mid 11 \mid AB \mid BA \mid BB \\ B & \to & 00 \end{array}$$

(g) Remove unit rule $A \to B$.

$$\begin{array}{lcl} S & \to & BAB \mid 00 \mid 11 \mid AB \mid BA \mid BB \mid \varepsilon \\ A & \to & BAB \mid 00 \mid 11 \mid AB \mid BA \mid BB \\ B & \to & 00 \end{array}$$

(h) Convert all rules to the proper form.

$$\begin{array}{lcl} S & \to & BA_1 \mid U_0U_0 \mid U_1U_1 \mid AB \mid BA \mid BB \mid \varepsilon \\ A & \to & BA_1 \mid U_0U_0 \mid U_1U_1 \mid AB \mid BA \mid BB \\ B & \to & U_0U_0 \\ A_1 & \to & AB \\ U_0 & \to & 0 \\ U_1 & \to & 1 \end{array}$$

Optimization was performed to reduce the number of variables in the conversion. $\square$

6. Draw the state diagram of a pushdown automaton (PDA) that recognizes the following language: $\{w\#x \mid w^{\mathcal{R}}$ is a substring of x for $w, x \in \{0,1\}^*\}$. Please explain the intuition behind the PDA.

Solution. To be provided. $\square$

7. Prove, using the pumping lemma, that $\{a^p \mid p$ is a prime number$\}$ is not context-free.

Solution. Suppose q is the pumping length. Consider a string $s = a^{q'}$, where q' is a prime number greater than or equal to q. We further suppose that s can be pumped by dividing s as $uvxyz = a^i a^j a^k a^l a^{q'-i-j-k-l}$, where $j + l > 0$ and $j + k + l \leq q \leq q'$.

We can pump s up to $a^i(a^j)^m a^k(a^l)^m a^{q'-i-j-k-l}$ for any $m > 1$, obtaining strings of the form $a^{jm+lm+q'-j-l} = a^{(j+l)(m-1)+q'}$. However, for $m = q' + 1$, $a^{(j+l)(m-1)+q'} = a^{(j+l)(q'+1-1)+q'} = a^{(j+l+1)q'}$ is clearly not in the language $\{a^p \mid p \text{ is a prime number}\}$. Thus, s cannot be pumped and the language is not context-free. $\square$

8. Prove that the language over $\{a, b, c\}$ with equal numbers of a's, b's, and c's is not context-free.

   Solution. Let A denote the language as defined in the problem statement. Let $B = \{a^i b^j c^k \mid i, j, k \geq 0\}$, which is apparently regular. As the intersection of a context-free language and a regular language is context-free (asserted in the appendix), if A were context-free, then $A \cap B$ would also be context-free. However, $A \cap B$ equals $\{a^n b^n c^n \mid n \geq 0\}$, which is not context-free (also asserted in the appendix), and it follows that A is not context-free. $\square$

9. Find a regular language A, a non-regular but context-free language B, and a non-context-free language C over $\{a, b\}$ such that $C \subseteq B \subseteq A$.

   Solution. $A = \{a^i b^j a^k \mid i, j, k \geq 0\}$ is regular. $B = \{a^i b^j a^k \mid i, j, k \geq 0 \text{ and } i \leq j\}$ is context-free but not regular. $C = \{a^i b^j a^k \mid i, j, k \geq 0 \text{ and } i \leq j \leq k\}$ is not context-free. It is apparent that $C \subseteq B \subseteq A$. $\square$

Appendix

- Properties of a binary relation R:
  - R is *reflexive* if for every x, xRx.
  - R is *symmetric* if for every x and y, xRy if and only if yRx.
  - R is *transitive* if for every x, y, and z, xRy and yRz implies xRz.
- A context-free grammar is in **Chomsky normal form** if every rule is of the form

$$\begin{array}{lcl} A & \rightarrow & BC \quad \text{or} \\ A & \rightarrow & a \end{array}$$

  where a is any terminal and A, B, and C are any variables—except that B and C may not be the start variable. In addition,

$$S \rightarrow \varepsilon$$

  is permitted if S is the start variable.
- If A is a regular language, then there is a number p (the pumping length) such that, if s is any string in A and $|s| \geq p$, then s may be divided into three pieces, $s = xyz$, satisfying the conditions: (1) for each $i \geq 0$, $xy^iz \in A$, (2) $|y| > 0$, and (3) $|xy| \leq p$.

- If A is a context-free language, then there is a number p such that, if s is a string in A and $|s| \geq p$, then s may be divided into five pieces, $s = uvxyz$, satisfying the conditions: (1) for each $i \geq 0$, $uv^ixy^iz \in A$, (2) $|vy| > 0$, and (3) $|vxy| \leq p$.
- The language $\{a^nb^nc^n \mid n \geq 0\}$ is not context-free.
- The intersection of a context-free language and a regular language is context-free.