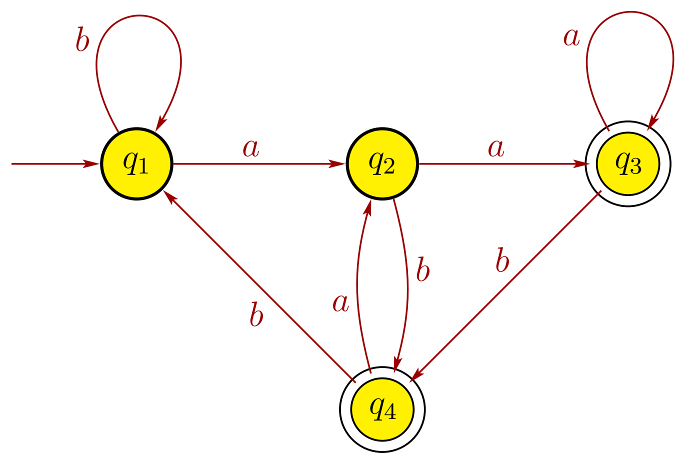
CS 341, Fall 2020, Hybrid Section
Solutions for Midterm 1

1. (a) False. If A is nonregular, then A cannot have an NFA by Corollary 1.40.
   (b) False. Let $A = \{ a^n b^n \mid n \geq 0 \}$ and $B = (a \cup b)^*$. Then $A \subseteq B$, A is nonregular, and B is regular.
   (c) False. Let $A = \emptyset$ and $B = \{ a^n b^n \mid n \geq 0 \}$. Then $A \subseteq B$, A is regular because it's finite, and B is nonregular.
   (d) False. The language a^* is regular but infinite.
   (e) False. $A = \{ a^n b^n \mid n \geq 0 \}$ is context-free but not regular.
   (f) False. Homework 6, problem 2b.
   (g) False. $0^*1^*0^*$ generates the string $001000 \notin A$, so the regular expression is not correct. In fact, A is nonregular, so it can't have a regular expression.
   (h) True. Theorem 1.54 implies A is regular. Then by Corollary 2.32, A is context-free, so Theorem 2.20 ensures that A has a PDA.
   (i) True, by Lemma 2.27 and Theorem 2.9.
   (j) False. The transition function of an NFA is $\delta : Q \times \Sigma_\varepsilon \to \mathcal{P}(Q)$.

2. (a) $b^*ab^* \cup b^*ba^*bb^*$. Another regular expression is $b^*(a \cup ba^*b)b^*$. There are infinitely many correct regular expressions for the language.
   (b) $G_3 = (V_3, \Sigma, R_3, S_3)$ with $S_3 \notin V_1 \cup V_2$, where
      - $V_3 = V_1 \cup V_2 \cup \{S_3\}$,
      - S_3 is the (new) starting variable,
      - Σ is the same alphabet of terminals as in G_1 and G_2, and
      - $R_3 = R_1 \cup R_2 \cup \{S_3 \to S_1 S_2\}$.
   (c) $M_3 = (Q_3, \Sigma, \delta_3, q_3, F_3)$, where
      - $Q_3 = Q_1 \times Q_2$;
      - Σ is the same alphabet as M_1 and M_2 have;
      - the transition function δ_3 satisfies $\delta_3((q,r),\ell) = (\delta_1(q,\ell), \delta_2(r,\ell))$ for $(q,r) \in Q_3$ and $\ell \in \Sigma$;
      - the starting state $q_3 = (q_1, q_2)$; and
      - $F_3 = (Q_1 \times F_2) \cap (F_1 \times Q_2)$, which also can be written as $F_1 \times F_2$.
   (d) After the one step of removing $A \to \varepsilon$, the CFG is then

$$
\begin{aligned}
S_0 &\to S \\
S &\to 0A1AS \mid 01AS \mid 0A1S \mid 01S \mid 0A1 \mid 01 \mid \varepsilon \\
A &\to 10A1 \mid 101
\end{aligned}
$$

3. (a) A DFA for A is below:



A 5-tuple description of the DFA above is $M = (Q, \Sigma, \delta, q_1, F)$, where

- $Q = \{q_1, q_2, q_3, q_4\}$
- $\Sigma = \{a, b\}$
- The transition function $\delta : Q \times \Sigma \to Q$ is defined as

	a	b
q_1	q_2	q_1
q_2	q_3	q_4
q_3	q_3	q_4
q_4	q_2	q_1

- q_1 is the start state
- $F = \{q_3, q_4\}$

There are infinitely many other correct DFAs for A.

(b) A regular expression for A is $(a \cup b)^* a (a \cup b)$. There are infinitely many other correct regular expressions for A.

4. $G = (V, \Sigma, R, S)$ with set of variables $V = \{S, Y\}$, where S is the start variable; set of terminals $\Sigma = \{a, b, c\}$; and rules

$$\begin{aligned} S &\to cSb \mid Y \\ Y &\to cYa \mid \varepsilon \end{aligned}$$

There are infinitely many other correct CFGs for D.

5. Language D is nonregular. We prove this by contradiction. Suppose that D is a regular language. Let p be the "pumping length" of the Pumping Lemma. Consider the string $s = c^p b^p$. Note that $s \in D$ because the number of c's equals the sum of the number of a's and the number of b's, and $|s| = 2p > p$, so the Pumping Lemma will hold. Thus, there exists strings x, y, and z such that $s = xyz$ and

(a) $xy^iz \in D$ for each $i \geq 0$,

(b) $|y| > 0$,

(c) $|xy| \leq p$.

Since the first p symbols of s are all c's, the third property implies that x and y consist only of c's. So z will be the rest of the c's, followed by b^p. The second property states that $|y| > 0$, so y has at least one c. More precisely, we can then say that

$$\begin{aligned} x &= c^j \text{ for some } j \geq 0, \\ y &= c^k \text{ for some } k \geq 1, \\ z &= c^m b^p \text{ for some } m \geq 0. \end{aligned}$$

Since $c^p b^p = s = xyz = c^j c^k c^m a^p = c^{j+k+m} b^p$, we must have that

$$j + k + m = p \quad \text{and} \quad k \geq 1.$$

The first property implies that $xy^2z \in D$, but

$$\begin{aligned} xy^2z &= c^j c^k c^k c^m b^p \\ &= c^{p+k} b^p \notin D \end{aligned}$$

since $p + k > p$ because $j + k + m = p$ and $k \geq 1$, so the number of c's in the pumped string xy^2z doesn't match the the number of a's and b's. Because the pumped string $xy^2z \notin D$, we have a contradiction. Therefore, D is a nonregular language.