CS 341-006, Spring 2021, Hybrid Section
Solutions for Midterm 1

1. (a) False. The language $A = \{ 0^n1^n \mid n \geq \}$ is context-free, but is nonregular, so A does not have an NFA.

   (b) False. Let A have the regular expression $(0 \cup 1)^*$, and let $B = \{ 0^n1^n \mid n \geq \}$. Then A is regular, B is nonregular, and $A \cup B = A$, which is regular.

   (c) False. Let $A = \emptyset$, and let $B = \{ 0^n1^n \mid n \geq \}$. Then A is regular, B is nonregular, and $A \cap B = A$, which is regular.

   (d) False. The language a^* is regular but infinite.

   (e) True. Let A be nonregular, and suppose for contradiction that $\overline{A}$ is regular. Because the class of regular languages is closed under complements, we must then have that the complement of $\overline{A}$ is regular. But the complement of $\overline{A}$ is $\overline{\overline{A}} = A$, which we said was nonregular, so we get a contradiction. Thus, $\overline{A}$ must be nonregular.

   (f) False. HW 6, problem 2(a).

   (g) False. The language A is non-context-free, which can be proven using the same basic proof on slides 2-96 and 2-97, so A cannot have a CFG.

   (h) True. If A has a regular expression, then A is a regular language by Kleene's Theorem. All regular languages are also context-free, so A must then be context-free, and A then has a PDA by Theorem 2.20.

   (i) False. The language $A = \{ a^nb^nc^n \mid n \geq 0 \}$ is nonregular. But A is also non-context-free (slides 2-96 and 2-97), so A cannot have a context-free grammar.

   (j) True. HW 5, problem 3b.

2. (a) $b^*(ba^*b \cup a)ab^*$. Other regular expressions for the language include $b^*ba^*bab^* \cup b^*aab^*$ and $b^*(ba^*b \cup a)ab^* \cup \emptyset$. 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
   - $R_3 = R_1 \cup R_2 \cup \{S_2 \to S_1 \mid S_2\}$.

   (c) $M_3 = (Q_3, \Sigma, \delta_3, q_3, F_3)$, where

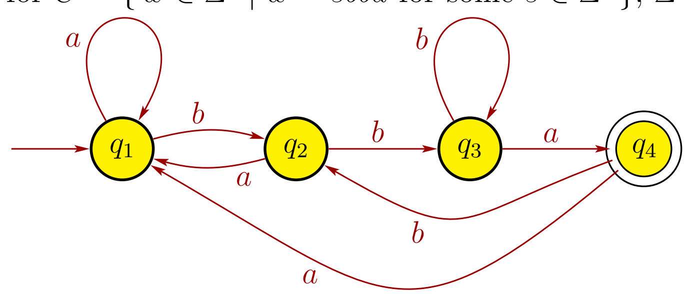
   - $Q_3 = Q_1 \times Q_2$;
   - Σ is the same alphabet as M_1 and M_2 have;
   - the transition function δ_3 satisfies $\delta_3((q, r), \ell) = (\delta_1(q, \ell), \delta_2(r, \ell))$ for $(q, r) \in Q_3$ and $\ell \in \Sigma$;
   - the starting state $q_3 = (q_1, q_2)$; and

- $F_3 = F_1 \times F_2$

(d) After the one step of removing $S \to \varepsilon$, the CFG is then

$$\begin{aligned} S_0 &\to S \mid \varepsilon \\ S &\to 0A1SA \mid 0A1A \\ A &\to 0S0 \mid 00 \mid A0S10S1 \mid A010S1 \mid A0S101 \mid A0101 \mid \varepsilon \end{aligned}$$

3. (a) A DFA for $C = \{ w \in \Sigma^* \mid w = sbba \text{ for some } s \in \Sigma^* \}$, $\Sigma = \{a, b\}$, 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_1	q_2
q_2	q_1	q_3
q_3	q_4	q_3
q_4	q_1	q_2

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

There are infinitely many other correct DFAs for C.

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

4. A CFG for $D = \{ a^i b^j \mid i \leq j \}$ is $G = (V, \Sigma, R, S)$ with set of variables $V = \{S, X\}$, where S is the start variable; set of terminals $\Sigma = \{a, b\}$; and rules

$$\begin{aligned} S &\to aSb \mid X \\ X &\to Xb \mid \varepsilon \end{aligned}$$

There are infinitely many other correct CFGs for D. For example, we could define R to instead be

$$\begin{aligned} S &\to aSb \mid X \\ X &\to bX \mid \varepsilon \end{aligned}$$

5. Language $E = \{ w \in \Sigma^* \mid w = w^{\mathcal{R}} \}$ with $\Sigma = \{0, 1\}$ is nonregular. We prove this by contradiction. Suppose that E is a regular language. Let p be the "pumping length" of the Pumping Lemma. Consider the string

$$s = a^p b a^p.$$

Note that $s \in E$ because $s^{\mathcal{R}} = s$. Also, the length of s is $|s| = 2p + 1 > p$, so the Pumping Lemma will hold. Thus, there exists strings x, y, and z such that $s = xyz$ and

(i) $xy^i z \in E$ for each $i \geq 0$,

(ii) $|y| > 0$,

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

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

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

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

$$j + k + m = p, \text{ where } k \geq 1$$

by (ii). The first property implies that $xy^2 z \in E$, but

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

because $(a^{p+k} b a^p)^{\mathcal{R}} = a^b a^{p+k} \neq a^b a^p$. Because the pumped string $xy^2 z \notin E$, we have a contradiction. Therefore, E is a nonregular language.

A string that will not work for getting a contradiction is $s = 0^p \in E$, which has $|s| \geq p$, so the pumping lemma will apply. Then we could let $x = z = \varepsilon$ and $y = 0^p$, and every pumped string $xy^i z = 0^{ip} \in E$, so there is no contradiction. There are many other strings that won't work.