CS 341, Fall 2011
Solutions for Midterm 1

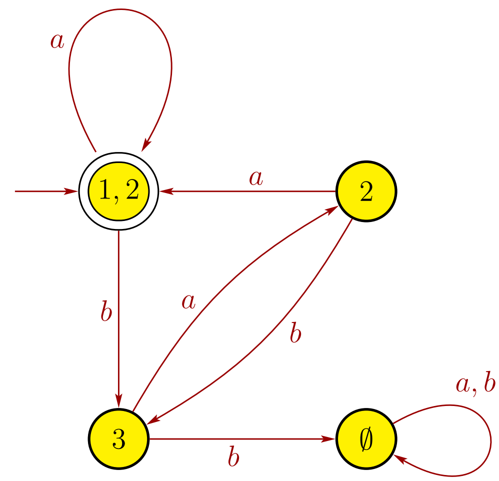
1. (a) False. The language a^* is regular but infinite.
   (b) True. Since A is finite (it has only 54 strings),
   (c) False. $\{a^n b^n c^n \mid n \geq 0\}$ is non-regular, but not context-free.
   (d) False. HW 6, problem 2(b).
   (e) True. By HW 2, problem 3, we know $\overline{A}$ is regular. Kleene's Theorem ensures B^* is regular, so $\overline{A} \cup B^*$ is regular by Theorem 1.25.
   (f) True. Corollary 2.32 implies A is context-free. Thus, A has a PDA by Theorem 2.20.
   (g) False. For example, $A = \{0^n 1^n \mid n \geq 0\}$ is a subset of $B = L((0 \cup 1)^*)$, but A is non-regular and B is regular.
   (h) False. Theorem 1.39.
   (i) True. By Theorem 2.9. The fact that A is non-regular is irrelevant.
   (j) True. Suppose A is non-context-free but regular. But then Corollary 2.32 implies A is context-free, which is a contradiction.

2. (a) $(a \cup ba^*b)(a \cup b)^*$
   (b) $(a \cup b)^*(ab \cup ba) \cup a \cup b \cup \varepsilon$
   (c) see slide 1-50.
   (d) CFG $G = (V, \Sigma, R, S)$ is Chomsky normal form means that each rule in R has one of 3 forms:

$$\begin{aligned} A &\rightarrow BC \\ A &\rightarrow x \\ S &\rightarrow \varepsilon \end{aligned}$$

   where $A \in V$; $B, C \in V - \{S\}$; $x \in \Sigma$, and S is the start variable. Thus, the violating rules are
   - $S \rightarrow aX$, since there cannot be a mix of terminals and variables on the RHS
   - $X \rightarrow Y$, since unit rules are not allowed
   - $Y \rightarrow \varepsilon$, since cannot have ε on RHS of rule when LHS is not S
   - $Y \rightarrow SX$, since S cannot be on RHS.

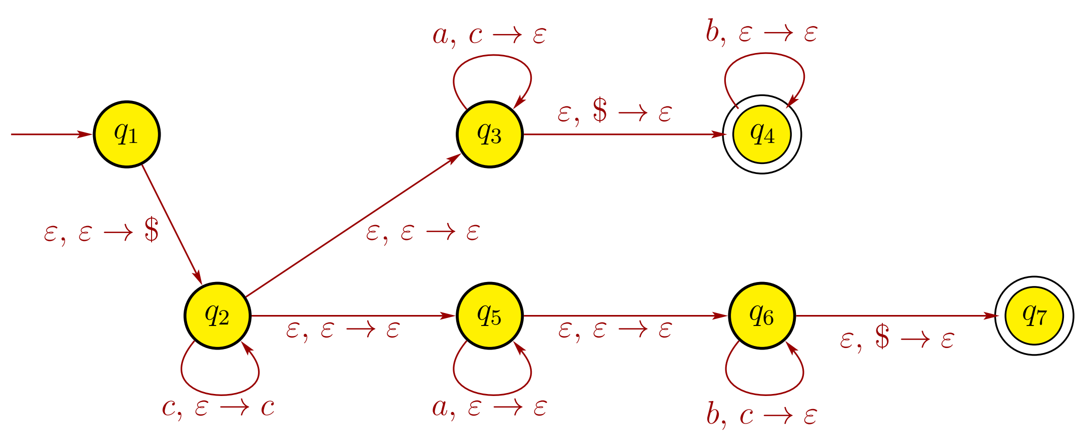
3. (a) $\varepsilon, a, aa, aaa, baa, \ldots$
   (b) A DFA for C is below:



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

$$\begin{aligned} S &\rightarrow UW \mid X \\ U &\rightarrow cUa \mid \varepsilon \\ W &\rightarrow bW \mid \varepsilon \\ X &\rightarrow cXb \mid Y \\ Y &\rightarrow aY \mid \varepsilon \end{aligned}$$

   (b) PDA



5. Language A is nonregular. We prove this by contradiction. Suppose that A is a regular language. Let p be the "pumping length" of the Pumping Lemma. Consider the string $s = a^p b a^p b a^p b$. Note that $s \in A$, and $|s| = 3p + 3 > p$, so the Pumping Lemma will hold. Thus, there exists strings x, y, and z such that $s = xyz$ and

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

   (b) $|y| > 0$,

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

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, followed by ba^pba^pb. The second property states that $|y| > 0$, so y has at least one a. More precisely, we can then say that

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

Since $a^pba^pba^pb = s = xyz = a^ja^ka^mba^pba^pb = a^{j+k+m}ba^pba^pb$, we must have that

$$j + k + m = p.$$

The first property implies that $xy^2z \in A$, and

$$\begin{aligned} xy^2z &= a^ja^ka^ka^mba^pba^pb \\ &= a^{p+k}ba^pba^pb \end{aligned}$$

since $j + k + m = p$. When xy^2z is split into equal thirds (assuming that this is even possible), the first third contains only a's since $k \geq 1$, but at least one of the other thirds contains a b. Hence, $xy^2z \notin A$, so we get a contradiction. Therefore, A is a nonregular language.