CS 341, Spring 2010
Solutions for Midterm 1

1. (a) 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.
   (b) False. Let $A = \emptyset$ and $B = \{ a^n b^n \mid n \geq 0 \}$. Then $A \subseteq B$, A is regular since it's finite, and ZB is nonregular.
   (c) False. The language a^* is regular but infinite.
   (d) False. $A = \{a^n b^n \mid n \geq 0\}$ is context-free but not regular.
   (e) True. Homework 2, problem 5.
   (f) False. 0^*1^* generate the string $001 \notin A$, so the regular expression is not correct. In fact, A is nonregular, so it can't have a regular expression.
   (g) False. If A has an NFA, then Corollary 1.40 implies that A is regular.
   (h) True. Corollary 2.32.
   (i) True, by Lemma 2.27 and Theorem 2.9.
   (j) False. The transition function of an NFA is $\delta : Q \times \Sigma_\varepsilon \to \mathcal{P}(Q)$.

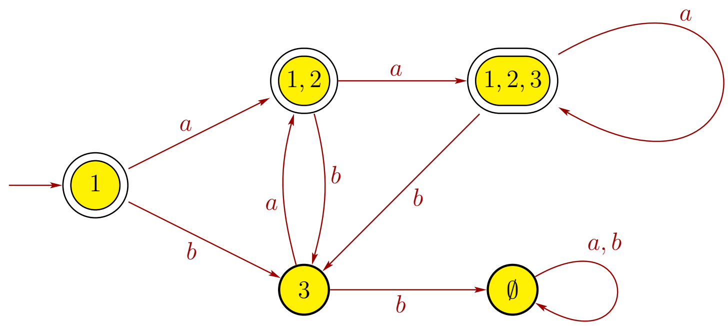
2. (a) $(a \cup bb^*a)(a \cup b)^*$
   (b) See slide 1-53.
   (c) CFG $G = (V, \Sigma, R, S)$ is Chomsky normal form means that each rule in R has one of 3 forms:

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

   where $A \in V$; $B, C \in V - \{S\}$; $x \in \Sigma$, and S is the start variable. Thus, the violating rules are

   - $S \to ba$, since there cannot be two terminals on the RHS
   - $S \to Yb$, since there cannot be a mix of variables and terminals on the RHS
   - $X \to YS$, since S cannot be on RHS
   - $X \to \varepsilon$, since cannot have ε on RHS of rule when LHS is not S
   - $Y \to X$, since it is a unit rule.

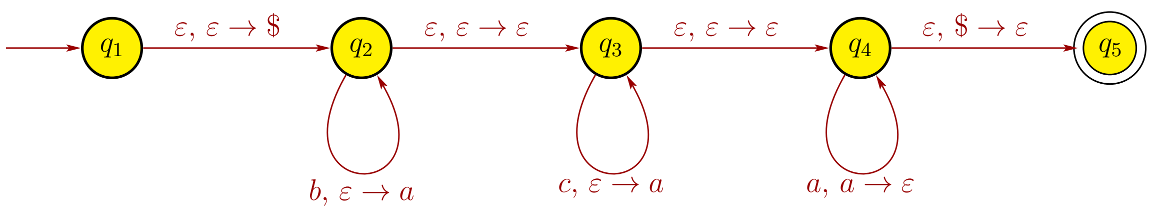
3. (a) $\varepsilon, a, aa, ba, aaa, \ldots$
   (b) A DFA for C is below:



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

$$\begin{aligned} S &\rightarrow bSa \mid X \\ X &\rightarrow cXa \mid \varepsilon \end{aligned}$$

   (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 $s = b^p c^p a^{2p}$. Note that $s \in A$, and $|s| = 4p > p$, 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 b's, the third property implies that x and y consist only of b's. So z will be the rest of the b's, followed by $c^p a^{2p}$. The second property states that $|y| > 0$, so y has at least one b. More precisely, we can then say that

$$\begin{aligned} x &= b^j \text{ for some } j \geq 0, \\ y &= b^k \text{ for some } k \geq 1, \\ z &= b^m c^p a^{2p} \text{ for some } m \geq 0. \end{aligned}$$

Since $b^p c^p a^{2p} = s = xyz = b^j b^k b^m c^p a^{2p} = b^{j+k+m} c^p a^{2p}$, we must have that

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

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

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

since $j+k+m = p$. Hence, $xy^2z \notin A$ since $k \geq 1$ so the number of b's plus the number of c's does not equal the number of a's, and we get a contradiction. Therefore, A is a nonregular language.