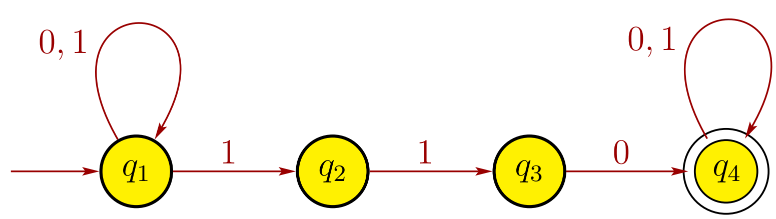
CS 341, Fall 2007
Solutions for Midterm 1

1. (a) False. If A has an NFA, then it is regular by Corollary 1.40.
   (b) True.
   (c) True. If A has a regular expression, then A is a regular language by Kleene's Theorem. Corollary 2.32 implies that A is CFL, so A has a CFG by definition.
   (d) False. The language $A = \{0^n1^n \mid n \geq 0\}$ has a PDA (see slide 2-54), but is not regular (slide 1-90), so A cannot have a DFA.
   (e) False. The language $A = \{0^n1^n \mid n \geq 0\}$ has a PDA (see slide 2-54), but is not regular (slide 1-90), so A cannot have a DFA. Thus, by Theorem 1.40, A cannot have an NFA either.
   (f) False. Homework 6, problem 2(a).
   (g) False. Homework 6, problem 2(b).
   (h) False. Every CFL has a CFG in Chomsky normal form by Theorem 2.9. But not every CFL is regular, e.g., $\{a^nb^n|n \geq 0\}$.
   (i) True, by Theorem 1.49.
   (j) True. If A_1 and A_2 are regular, then $A_1 \circ A_2$ is regular by Theorem 1.47. Corollary 2.32 then implies that $A_1 \circ A_2$ is context-free.

2. (a) CFG $G = (V, \Sigma, R, S)$ is Chomsky normal form means that each rule in R has one of 3 forms:

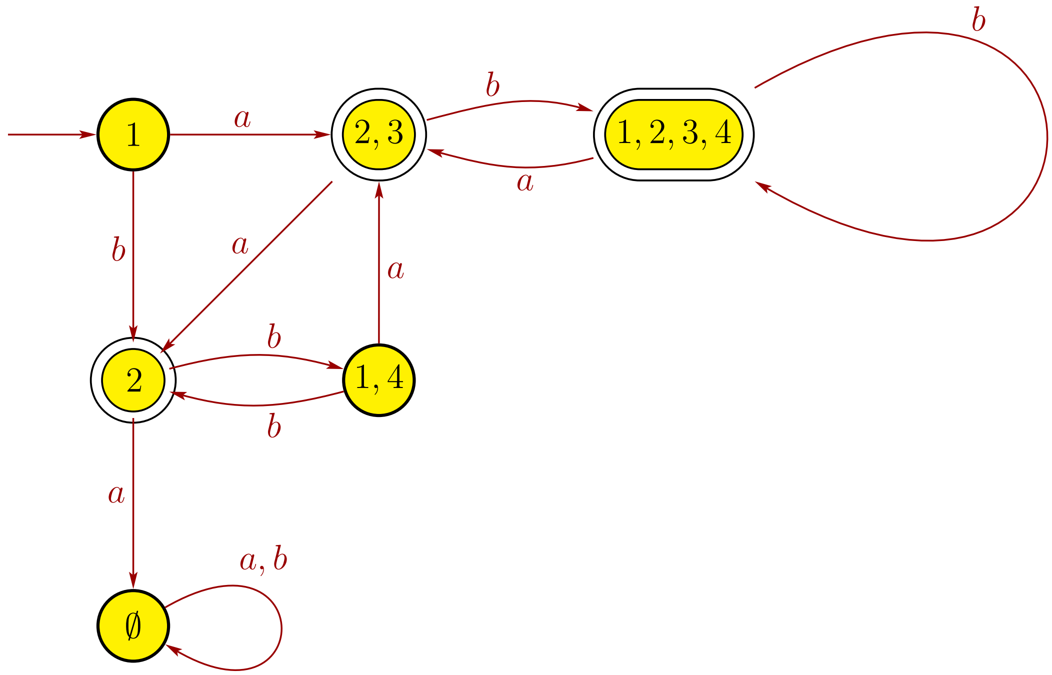
$$\begin{aligned} A &\rightarrow BC \\ A &\rightarrow x \\ S &\rightarrow \varepsilon \end{aligned}$$

   where $A \in V$; $B, C \in V - \{S\}$; $x \in \Sigma$, and S is the start variable.
   (b) Here is an NFA with exactly 4 states:



   (c) See page 1-53 of notes.
   (d) Homework 5, problem 3(a).

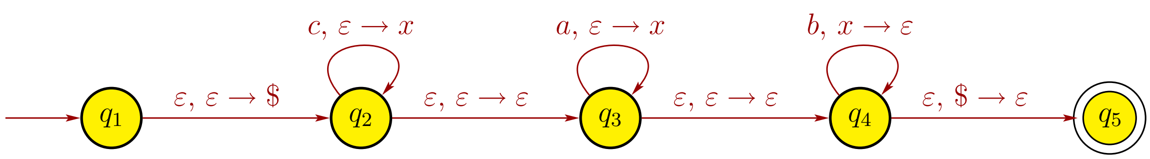
3. (a) $a, b, aa, ab, aba, \ldots$
   (b) Here's a DFA for C.



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

$$\begin{aligned} S &\rightarrow cSb \mid X \\ X &\rightarrow aXb \mid \varepsilon \end{aligned}$$

(b)



5. 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 $s = b^p a^{p+1}$. Note that $s \in A$, and $|s| = 2p + 1 > p$, 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 a^{p+1}. The second property states that $|y| > 0$, so y has at least one b. More precisely, we can then say that

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

Since $b^p a^{p+1} = s = xyz = b^j b^k b^m a^{p+1} = b^{j+k+m} a^{p+1}$, we must have that $j + k + m = p$. The first property implies that $xy^2z \in A$, but

$$\begin{aligned} xy^2z &= b^j b^k b^k b^m a^{p+1} \\ &= b^{p+k} a^{p+1} \end{aligned}$$

since $j + k + m = p$. Hence, $xy^2z \notin A$ because $k \geq 1$, and we get a contradiction. Therefore, A is a nonregular language.