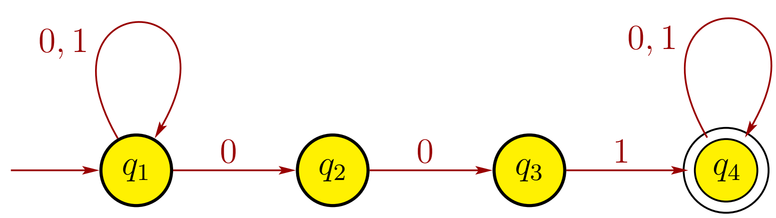
CS 341, Spring 2008
Solutions for Midterm 1

1. (a) False. If a language is nonregular, it cannot have an NFA by Corollary 1.40.
   (b) False. The language is non-regular, as shown on slide 1-90. Therefore, it cannot have a regular expression. The problem with the expression given is that it goes on forever, but regular expressions can't do that.
   (c) False. The language a^* is regular but infinite.
   (d) True. If A has a DFA, then it is regular. But every regular language is also context-free by Corollary 2.32. Therefore, A must have a PDA by Theorem 2.20.
   (e) False. The language $A = \{0^n 1^n 2^n \mid n \geq 0\}$ is non-context-free, so it cannot have a PDA by Theorem 2.20.
   (f) False. Homework 6, problem 2(a).
   (g) True. Homework 5, problem 3(b).
   (h) True. Kleene's Theorem.
   (i) True, Homework 2, problem 3.
   (j) True. Slide 1-10 gives a DFA for $\emptyset$, so it is regular.

2. (a) Here is an NFA with exactly 4 states:



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

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

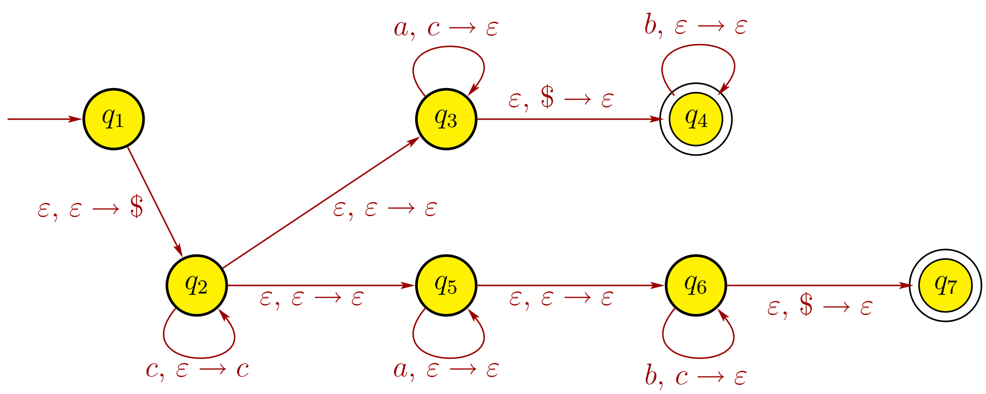
   where $A \in V$; $B, C \in V - \{S\}$; $x \in \Sigma$, and S is the start variable.
   (c) See page 1-50 of notes.
   (d) Homework 5, problem 3(c).

3. (a) $a, b, aa, ba, bb, \ldots$
   (b) A regular expression for C is $(a \cup ba^*(b \cup \varepsilon))(aa \cup aba^*(b \cup \varepsilon))^*(\varepsilon \cup a)$. There are infinitely many correct regular expressions for C.

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

$$\begin{aligned} S &\rightarrow X \mid Y \\ X &\rightarrow UT \\ U &\rightarrow cUa \mid \varepsilon \\ T &\rightarrow bT \mid \varepsilon \\ Y &\rightarrow cYb \mid Z \\ Z &\rightarrow aZ \mid \varepsilon \end{aligned}$$

(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 $s = (^p0)^p$, which is p left parentheses, followed by 0, followed by p right parentheses. Note that $s \in A$ since we can derive it as follows: $S \Rightarrow (S) \Rightarrow ((S)) \Rightarrow \cdots \Rightarrow (^pS)^p \Rightarrow (^p0)^p$. Also, $|s| = 2p + 1 > p$, so the Pumping Lemma will hold. Thus, there exists strings x, y, and z such that $s = xyz$ and

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

(b) $|y| > 0$,

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

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

$$\begin{aligned} x &= (^j \text{ for some } j \geq 0, \\ y &= (^k \text{ for some } k \geq 1, \\ z &= (^m0)^p \text{ for some } m \geq 0. \end{aligned}$$

Since $(^p 0)^p = s = xyz = (^j (^k (^m 0)^p = (^{j+k+m} 0)^p$, we must have that $j + k + m = p$. The first property implies that $xy^2z \in A$, but

$$\begin{aligned} xy^2z &= (^j (^k (^k (^m 0)^p \\ &= (^{p+k} 0)^p \end{aligned}$$

since $j + k + m = p$. Hence, $xy^2z \notin A$ because the parentheses don't balance since $k \geq 1$, and we get a contradiction. Therefore, A is a nonregular language.