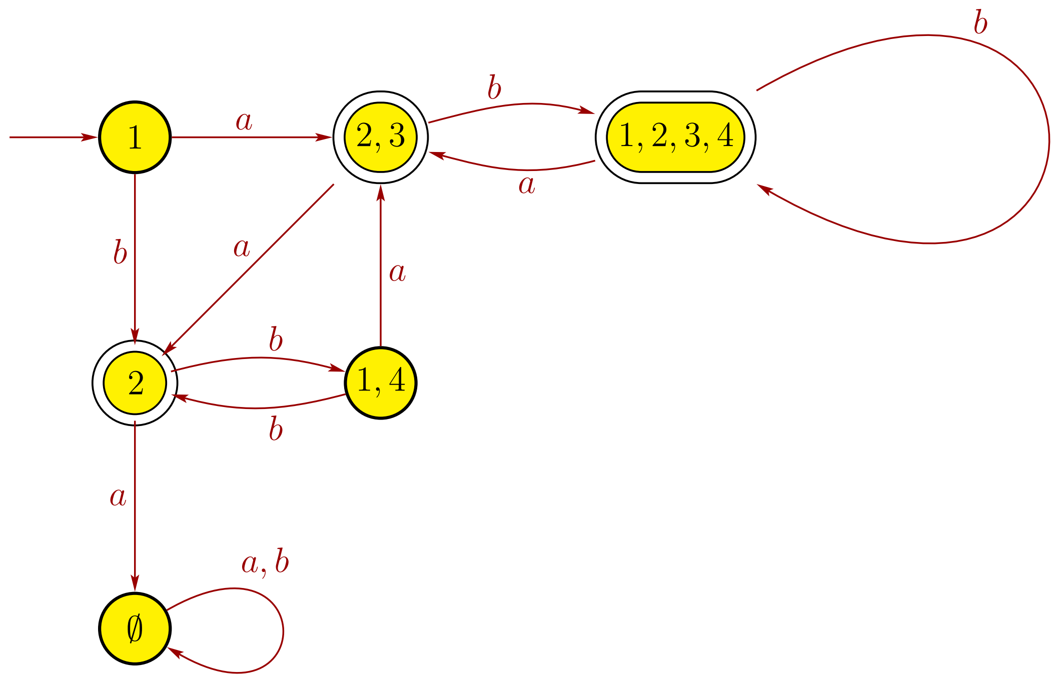
CS 341, Fall 2014, Face-to-Face Section
Solutions for Midterm 1

1. (a) True, by Theorem 1.49.
   (b) 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.
   (c) False. If A has an NFA, then it is regular by Corollary 1.40.
   (d) True.
   (e) 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.
   (f) False. The language $A = \{0^n 1^n \mid n \geq 0\}$ has a PDA (see slide 2-50), but is not regular (slide 1-90), so A cannot have a DFA.
   (g) False. The language $A = \{0^n 1^n \mid n \geq 0\}$ has a PDA (see slide 2-54), but is not regular (slide 1-105), so A cannot have a DFA. Thus, by Theorem 1.40, A cannot have an NFA either.
   (h) False. Homework 6, problem 2(a).
   (i) False. Homework 6, problem 2(b).
   (j) False. Every CFL has a CFG in Chomsky normal form by Theorem 2.9. But not every CFL is regular, e.g., $\{0^n 1^n | n \geq 0\}$.

2. (a) $b^*ab^* \cup b^*aa^*bb^*$. Another regular expression is $b^*(a \cup aa^*b)b^*$. There are infinitely many regular expressions for the language.
   (b) $G' = (V', \Sigma, R', S_0)$, where $V' = V \cup \{S_0\}$, S_0 is the (new) starting variable, Σ is the same alphabet of terminals as in G, and $R' = R \cup \{S_0 \to SS_0 \mid \varepsilon\}$.
   (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_2 satisfies $\delta((q,r),\ell) = (\delta_1(q,\ell), \delta_2(r,\ell))$ for $(q,r) \in Q_3$ and $\ell \in \Sigma$; the starting state $q_3 = (q_1, q_2)$; and $F_3 = (Q_1 \times F_2) \cap (F_1 \times Q_2)$, which also can be written as $F_1 \times F_2$.
   (d) After one step, the CFG is then

$$\begin{aligned} S_0 &\to S \\ S &\to A1SA \mid 1SA \mid A1S \mid 1S \mid A0 \mid 0 \mid \varepsilon \\ A &\to 0S0 \end{aligned}$$

3. A DFA for C is below:

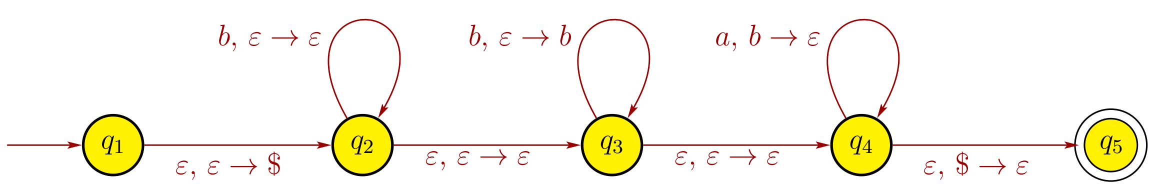


4. (a) $G = (V, \Sigma, R, S)$ with set of variables $V = \{S, Z\}$, where S is the start variable; set of terminals $\Sigma = \{a, b\}$; and rules

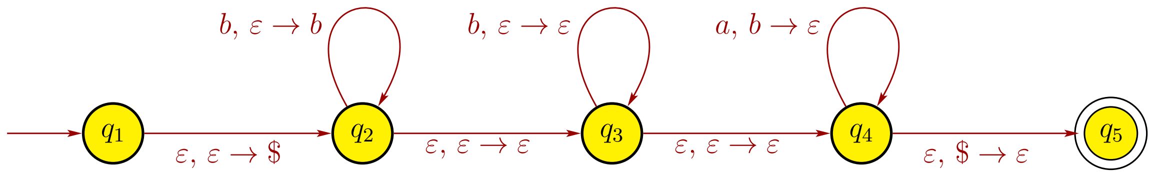
$$
\begin{aligned}
S &\rightarrow bSa \mid Z \\
Z &\rightarrow bZ \mid \varepsilon
\end{aligned}
$$

There are infinitely many other correct CFGs for L.

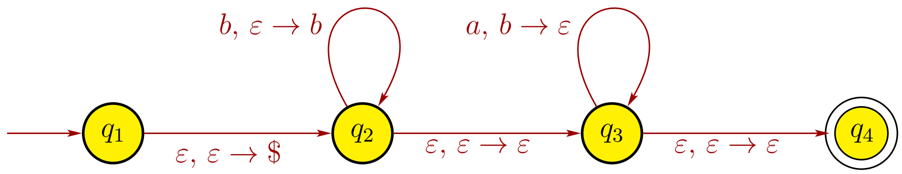
(b) There are infinitely many correct PDAs for L. The below PDA guesses how many b's not to match to the a's (state q_2), then pushes the b's to match with the a's (state q_3), matches the a's with the pushed b's (state q_4), and finally checks that the stack is empty (transition from q_4 to q_5).



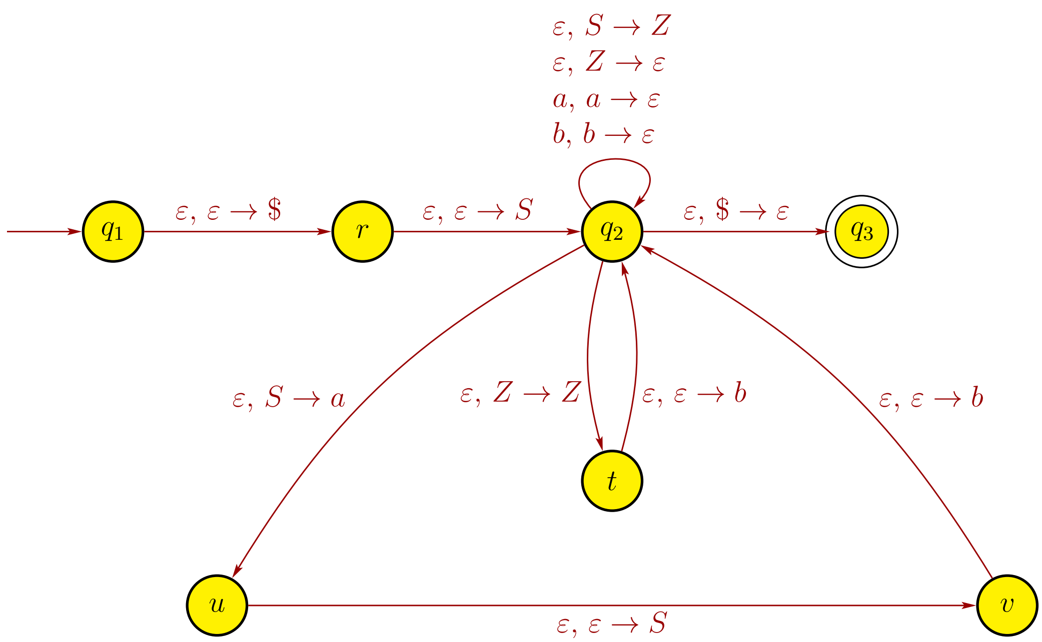
Below is another PDA for L, which first pushes b's to match the a's (state q_2), then guesses how many b's not to match with a's (state q_3), matches the a's with the pushed b's (state q_4), and finally checks that the stack is empty (transition from q_4 to q_5).



Below is yet another PDA for L. This one pushes all of the b's onto the stack (state q_2), and matches the a's with some of the pushed b's (state q_3). This PDA can accept a string with symbols (b's and \$) still on the stack.



Yet another approach uses the algorithm from Lemma 2.21 to convert the CFG in part (a) into a 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 $s = b^p a^p$. Note that $s \in A$, and $|s| = 2p > p$, 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. The second property states that $|y| > 0$, so y has at least one b. More precisely, we can then say that

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

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

$$j + k + m = p \quad \text{and} \quad k \geq 1.$$

The first property implies that $xy^0 z = xz \in A$, but

$$\begin{aligned} xz &= b^j b^m a^p \\ &= b^{j+m} a^p \notin A \end{aligned}$$

since $j + m < p$ because $j + k + m = p$ and $k \geq 1$, so the number of b's in s is less than the number of a's. This is a contradiction. Therefore, A is a nonregular language.