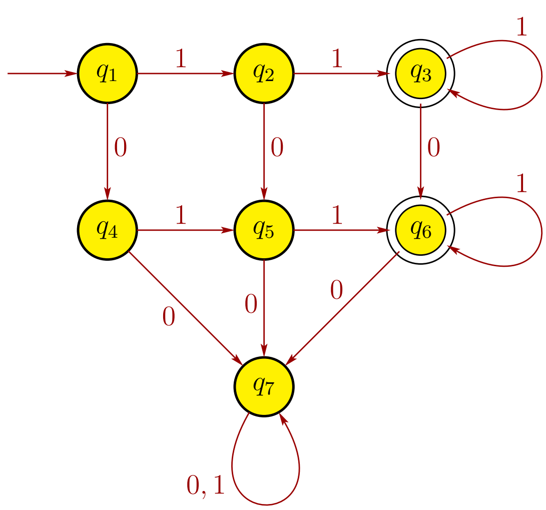
CS 341, Spring 2020, Hybrid Section
Solutions for Midterm 1

1. (a) False. $A = \{ a^n b^n \mid n \geq 0 \}$ is context-free but not regular.
   (b) True. Homework 2, problem 5.
   (c) False. $0^*1^*0^*$ generates the string $001000 \not\in A$, so the regular expression is not correct. In fact, A is nonregular, so it can't have a regular expression.
   (d) True. Theorem 1.54 implies A is regular. Then by Corollary 2.32, A is context-free, so Theorem 2.20 ensures that A has a PDA.
   (e) True, by Lemma 2.27 and Theorem 2.9.
   (f) False. The transition function of an NFA is $\delta : Q \times \Sigma_\varepsilon \to \mathcal{P}(Q)$.
   (g) False. If A is nonregular, then A cannot have an NFA by Corollary 1.40.
   (h) False. Let $A = \{ a^n b^n c^n \mid n \geq 0 \}$ and B be the language with regular expression $(a \cup b \cup c)^*$. Then $A \subseteq B$, A is non-context-free (slide 2-96), and B is context-free (because it is regular as it has a regular expression, so B is also context-free by Corollary 2.32).
   (i) False. Let $A = \emptyset$ and $B = \{ a^n b^n \mid n \geq 0 \}$. Then $A \subseteq B$, A is regular because it's finite, and B is nonregular.
   (j) False. The language a^* is regular but infinite.

2. (a) Here is one DFA for A. There are infinitely many other correct ones.

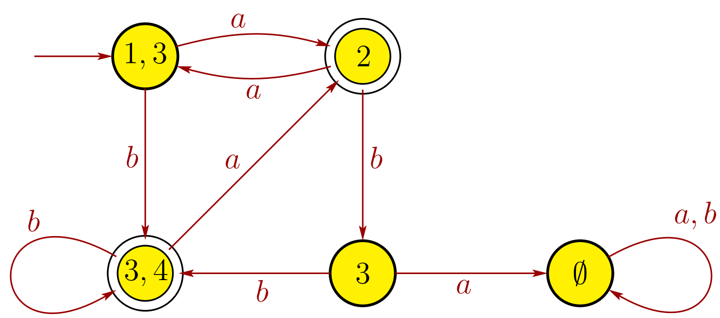


   (b) There are infinitely many correct regular expressions for the language. For example, $111^* \cup 111^*0 \cup 111^*01^* \cup 0111^* \cup 1011^*$ or $(11 \cup 111^*0 \cup 011 \cup 101)1^*$ or ....
   (c) $G' = (V', \Sigma, R', S_0)$, where $V' = V \cup \{S_0\}$, S_0 is the (new) starting variable, Σ is the same alphabet of terminals as in G, and $R' = R \cup \{S_0 \to SS_0 \mid \varepsilon\}$.

(d) After one step, the CFG is then

$$\begin{aligned} S_0 &\rightarrow S \\ S &\rightarrow 1SA0A \mid 1S0A \mid 1SA0 \mid 1S0 \mid 0AS1S \mid 0S1S \mid \varepsilon \\ A &\rightarrow 10S1 \end{aligned}$$

3. A DFA for C is below:

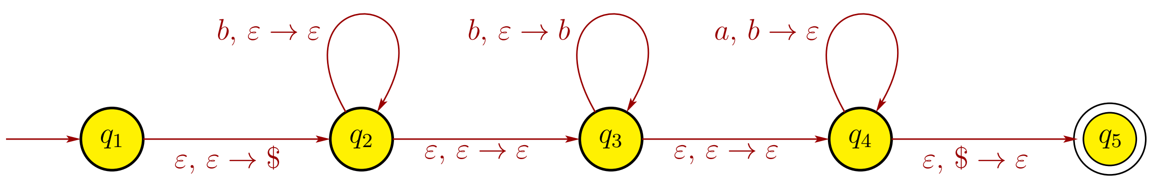


4. (a) $G = (V, \Sigma, R, S)$ with set of variables $V = \{S, Z\}$, where S is the start variable; set of terminals $\Sigma = \{a, b\}$; and rules

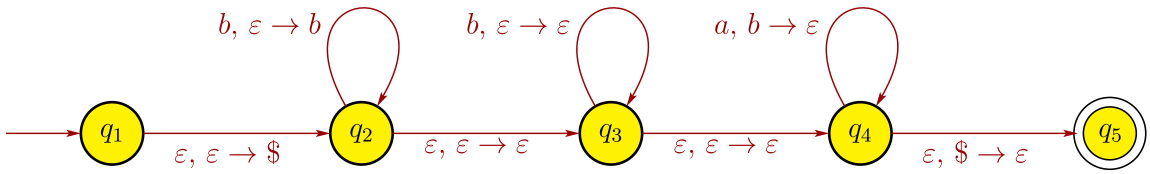
$$\begin{aligned} S &\rightarrow bSa \mid Z \\ Z &\rightarrow bZ \mid \varepsilon \end{aligned}$$

There are infinitely many other correct CFGs for L.

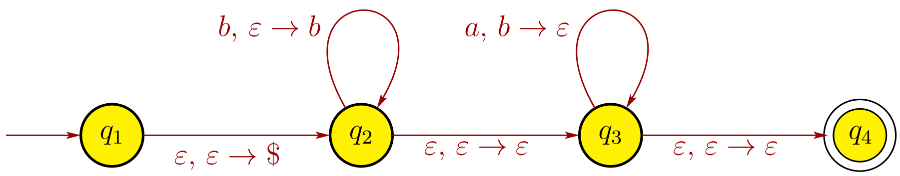
(b) There are infinitely many correct PDAs for L. The below PDA guesses how many b's not to match to the a's (state q_2), then pushes the remaining b's to match with the a's (state q_3), matches the a's with the pushed b's (state q_4), and finally checks that the stack is empty (transition from q_4 to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



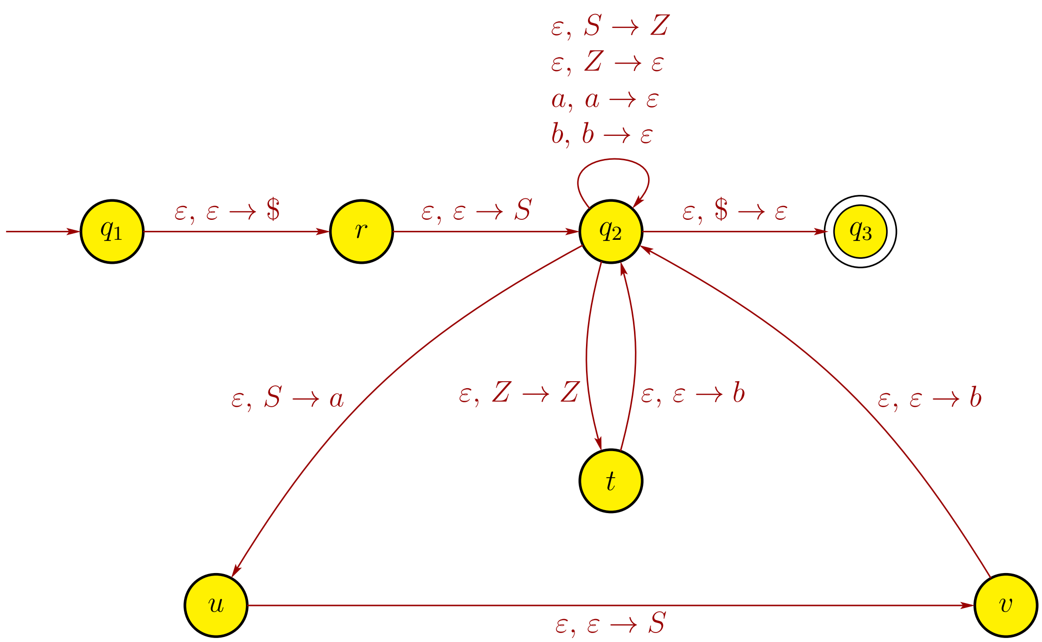
Below is another PDA for L, which first guesses how many b's to push to match the a's (state q_2), then reads (without pushing) the remaining b's (state q_3), matches the a's with the pushed b's (state q_4), and finally checks that the stack is empty (transition from q_4 to q_5).



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.