CS 341, Fall 2015, Face-to-Face Section
Solutions for Midterm 1

1. (a) False. 1^* is regular since it has a regular expression, but this language is infinite.

   (b) True. By HW 2, problem 3, we know $\overline{A}$ is regular. Since $\overline{A}$ and B are regular, then $\overline{A} \cup B$ is regular by Theorem 1.25. Theorem 1.49 then implies $(\overline{A} \cup B)^*$ is regular.

   (c) False. See HW 6, problem 2(a).

   (d) False. For example, $A = \{0^n 1^n 0^n \mid n \geq 0\}$ is a subset of $B = L((0 \cup 1)^*)$, but A is non-context-free and B is context-free.

   (e) True. Use the pumping lemma with string $a^r b^r$, where $r = \max(p, 3)$ and p is the pumping length.

   (f) True. Since A has a regular expression, A is a regular language by Theorem 1.54. Then Corollary 2.32 implies A is also context-free, so it has a CFG. Theorem 2.9 then ensures that A has a CFG in Chomsky normal form.

   (g) True. See slide 2-111.

   (h) True. HW 4, problem 5(a).

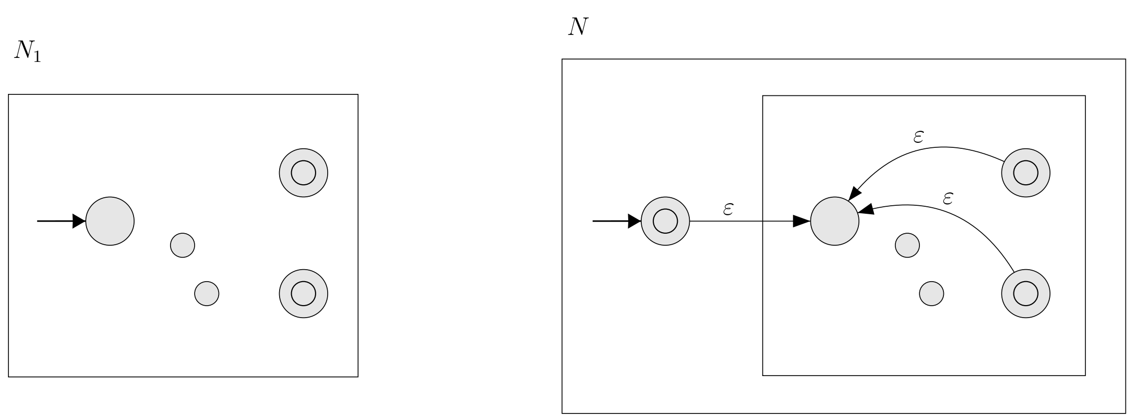
   (i) True. HW 4, problem 5(c).

   (j) True. Suppose A is non-context-free but regular. But then Corollary 2.32 implies A is context-free, which is a contradiction.

2. (a) $(\varepsilon \cup 1)(01)^*00(10)^*(\varepsilon \cup 1) \ \cup \ (\varepsilon \cup 0)(10)^*11(01)^*(\varepsilon \cup 0)$

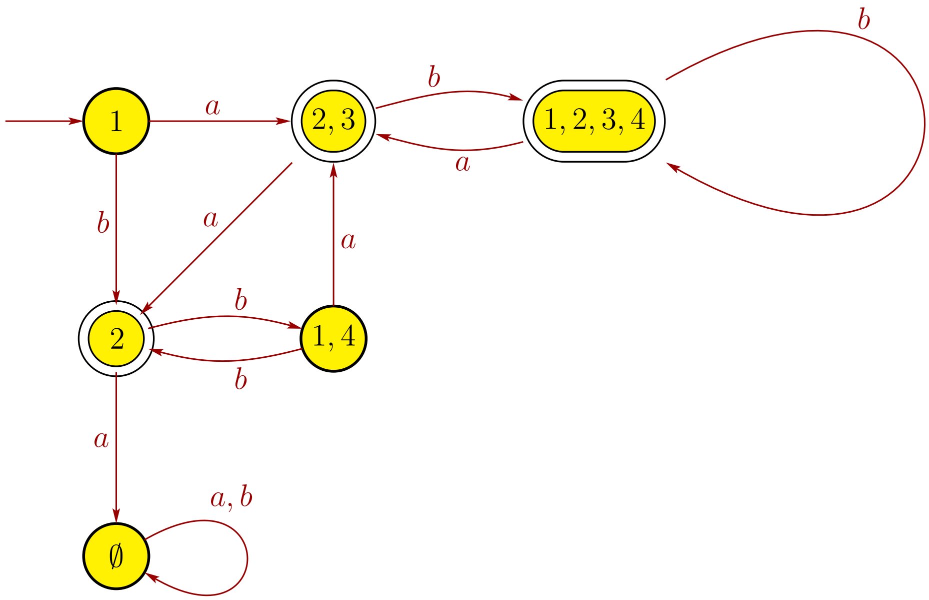
   (b) $(aa \cup b)a^*bb^*$ or $(aaa^* \cup ba^*)bb^*$ or ...

   (c) As on slide 1-66, an NFA N for A_1^* is as below:



   (d) (Homework 5, problem 3b.) Assume that $S_3 \notin V_1 \cup V_2$. Then a CFG for $A_1 \circ A_2$ is $G_3 = (V_3, \Sigma, R_3, S_3)$ with $V_3 = V_1 \cup V_2 \cup \{S_3\}$ and $R_3 = R_1 \cup R_2 \cup \{ S_3 \to S_1 S_2 \}$.

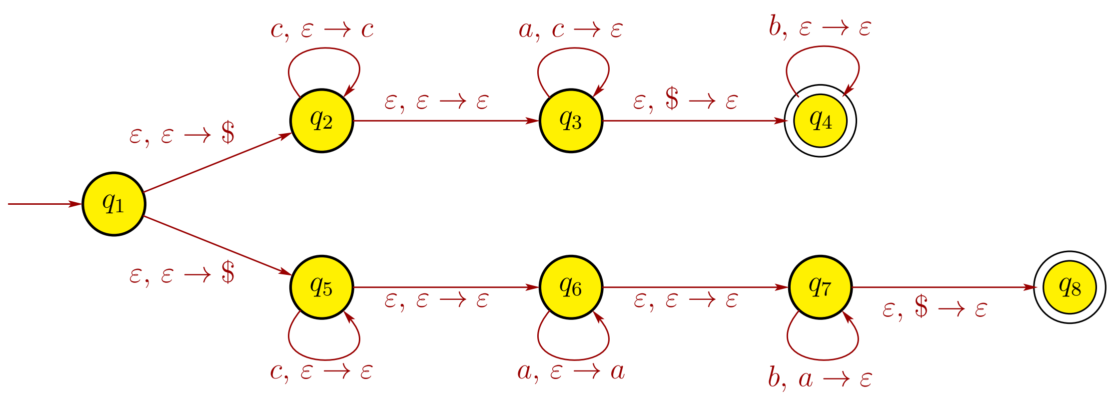
3. A DFA for C is below:



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

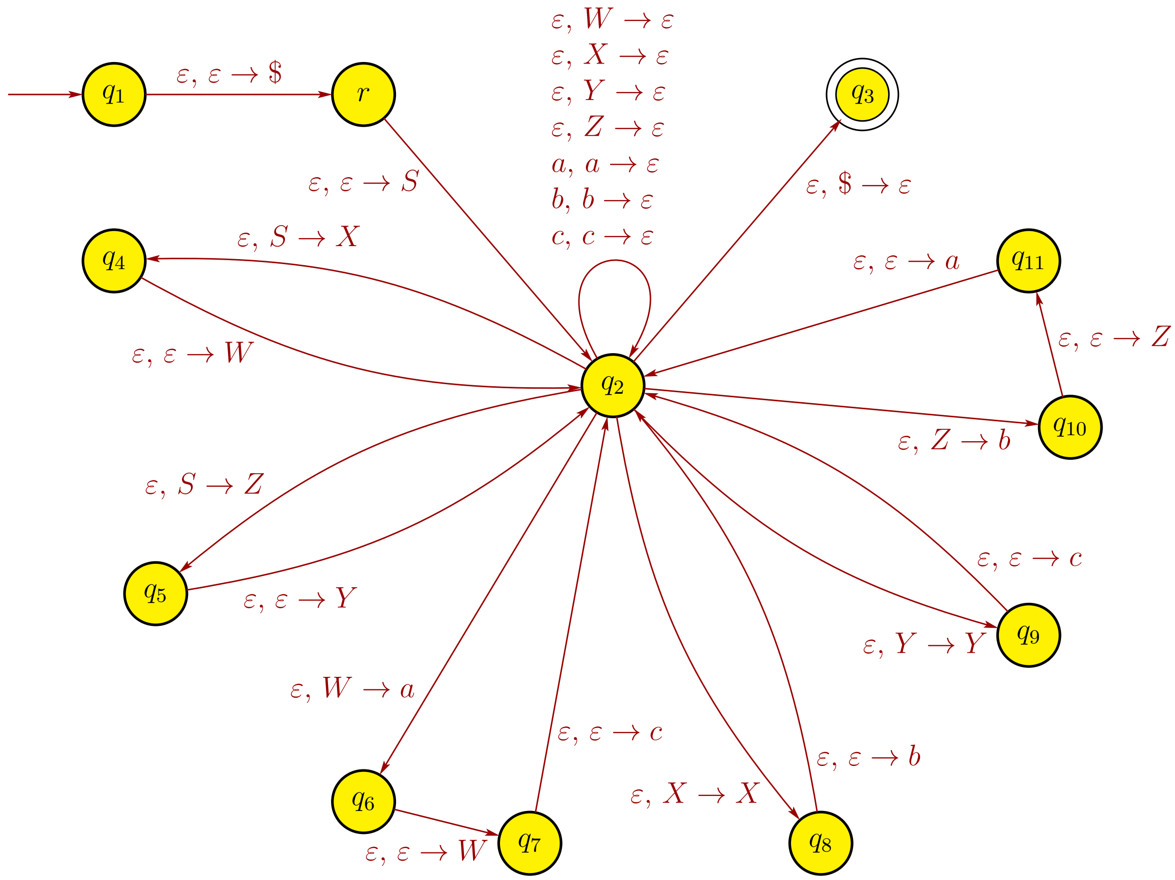
$$\begin{aligned} S &\rightarrow WX \mid YZ \\ W &\rightarrow cWa \mid \varepsilon \\ X &\rightarrow bX \mid \varepsilon \\ Y &\rightarrow cY \mid \varepsilon \\ Z &\rightarrow aZb \mid \varepsilon \end{aligned}$$

(b) There are infinitely many correct PDAs for L. Here is one:



The PDA has a nondeterministic branch at q_1. If the string is $c^i a^j b^k$ with $i = j$, then the PDA takes the branch from q_1 to q_2. If the string is $c^i a^j b^k$ with $j = k$, then the PDA takes the branch from q_1 to q_5.

Yet another approach uses the algorithm from Lemma 2.21 to convert the CFG in part (a) into a PDA.



Note that

- The path $q_2 \to q_4 \to q_2$ corresponds to the rule $S \to WX$.
- The path $q_2 \to q_5 \to q_2$ corresponds to the rule $S \to YZ$.
- The path $q_2 \to q_6 \to q_7 \to q_2$ corresponds to the rule $W \to cWa$.
- The path $q_2 \to q_8 \to q_2$ corresponds to the rule $X \to bX$.
- The path $q_2 \to q_9 \to q_2$ corresponds to the rule $Y \to cY$.
- The path $q_2 \to q_{10} \to q_{11} \to q_2$ corresponds to the rule $Z \to aZb$.

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 $s = c^p a^p$. Note that $s \in A$ because the numbers of c's and a's are equal, and $|s| = 2p > p$, so the Pumping Lemma will hold. Thus, there exists strings x, y, and z such that $s = xyz$ and

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

   (b) $|y| > 0$,

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

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

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

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

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

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

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

since $p + k > p$ because $j + k + m = p$ and $k \geq 1$, so the number of c's in the pumped string xy^2z doesn't match the number of a's, and the number of a's doesn't match the number of b's (none). Because the pumped string $xy^2z \notin A$, we have a contradiction. Therefore, A is a nonregular language.

Note that if you instead chose the string $s = c^p a^p b^p$, you would not get a contradiction. This is because pumping up or down leads to the number of c's changing, but the number of a's and b's remain the same and equal. Thus, the pumped string is still in the language, so there is no contradiction.