CS 341-008, Spring 2023, Face-to-Face Section
Solutions for Midterm 1

1. Multiple choice.

   1.1. Answer: (c).

   - The languages $L_1 = \{ a^n b^n c^n \mid n \geq 0 \}$ and $L_2 = \{ b^n a^n c^n \mid n \geq 0 \}$ are non-context-free languages (slide 2-96), with $L_1 \cap L_2 = \{\varepsilon\}$, which is regular because it is finite (slide 1-95). Thus, the intersection is also context-free by Corollary 2.32, making (a) incorrect.
   - If $L_1 = L_2 = \{ a^n b^n c^n \mid n \geq 0 \}$, then $L_1 \cap L_2 = L_1$, which is non-regular and non-context-free, so (b) and (d) are incorrect.
   - The previous two examples show that (c) is correct.

   1.2. Answer: (e).

   - Consider the language $A = \{ 0^n 1^n \mid n \geq 0 \}$, which we know is nonregular (slide 1-105). Now let $L = A^*$, which we can prove is also nonregular by the pumping lemma, which shows that (a) is incorrect. For an outline of the proof that L is nonregular, suppose that L is regular, and consider the string $s = 0^p 1^p \in L$, where p is the pumping length. Note that $|s| = 2p \geq p$, so the conclusions of the pumping lemma will hold. Thus, we can write $s = xyz$ with $x = 0^j$ for $j \geq 0$, $y = 0^k$ for $k \geq 1$, and $z = 0^m 1^p$ for $m \geq 0$, where $j + k + m = p$. But the pumped string $xyyz = 0^{p+k} 1^p$ cannot be written as a concatenation of zero of more strings from A. This contradicts the pumping lemma so L is nonregular, showing that (a) is incorrect. Also, let $B = A$, so $A \cap B = A$, which is nonregular, so (c) is also incorrect.
   - For the same language $A = \{ 0^n 1^n \mid n \geq 0 \}$, let $B = \{\varepsilon\}$, so $A \circ B = A$, which we know is nonregular. Thus, (b) is incorrect.

   1.3. Answer: (d).

   - Slide 2-99 proves that A is non-context-free, so (d) is correct.
   - Because A is non-context-free, Theorem 2.20 shows that A cannot have a PDA, making (c) incorrect.
   - Also, A being non-context-free implies that A is also not regular (Corollary 2.32), so (a) and (b) are incorrect. We can see that the regular expression $(0 \cup 1)^*(0 \cup 1)^*$ in (a) is wrong because it generates the string $01 \notin A$.

   1.4. Answer: (c).

   - HW 6, problem 2a, shows that the class of CFLs is not closed under intersection, so (a) is incorrect.
   - HW 6, problem 2b, shows that the class of CFLs is not closed under complementation, so (b) is incorrect.
   - HW 5, problem 3b, shows that (c) is correct.
   - The language $\{ a^n b^n c^n \mid n \geq 0 \}$ is not context-free by slide 2-96, so not all languages are context-free, so (d) is incorrect.

- By Corollary 2.32, every regular language is also context-free, so (e) is incorrect.

1.5. Answer: (c).

- The class of context-free languages is closed under union (Homework 5, problem 3a), so $B \cup C$ is context-free. Also, the class of context-free languages is closed under concatenation (Homework 5, problem 3b), ensuring that $A(B \cup C)$ is context-free, so (c) is correct.
- We know that the class of context-free languages is *not* closed under complementation (Homework 6, problem 2b), so there exists some context-free language D whose complement $\overline{D}$ is not context-free. Also, let $B = C = \{\varepsilon\}$, which is finite, so B and C are regular (slide 1-95), making them also context-free (Corollary 2.32). Thus, $B \cup C = \{\varepsilon\}$, and let $A = D$, so $\overline{A}(B \cup C) = \overline{A} = \overline{D}$ is non-context free, making (a) incorrect.
- Let A be any regular language, so A is also context-free (Corollary 2.32). As A is regular, $\overline{A}$ is also regular (Homework 2, problem 3), so $\overline{A}$ is also context-free (Corollary 2.32). The class of context free languages is closed under concatenation (Homework 5, problem 3b) and union (Homework 5, problem 3a), so in this case when A is regular, we have that $\overline{A}(B \cup C)$ is context-free, showing (b) is incorrect.
- For $\Sigma = \{a, b\}$, let A be the language of all strings over Σ that don't begin with a. Now A has regular expression $\varepsilon \cup b(a \cup b)^*$, so Kleene's Theorem implies that A is a regular language, making A also context-free (Corollary 2.32). Also, $\overline{A}$ is the set of all strings over Σ that begin with a; e.g., $a \in \overline{A}$. Also, let $B = \{b\}$ and $C = \{b\}$, each of which are finite so also regular (slide 1-95) and context-free (Corollary 2.32). Also, $B \cup C = \{b\}$. Then, we have that $ab \in \overline{A}(B \cup C)$, but $ab \notin (B \cup C)\overline{A}$, making (d) incorrect.

1.6. Answer: (h).

- The regular expression $(00 \cup 11)^*$ cannot generate the string $1010 \in L$, so (i) is incorrect.
- The regular expression $(0(00 \cup 11)^*0 \ \cup \ 1(00 \cup 11)^*1)^*$ cannot generate the string $0101 \in L$, so (ii) is incorrect.
- The regular expression $(00 \cup 11 \cup 0(00 \cup 11)^*0 \cup 1(00 \cup 11)^*1)^*$ cannot generate the string $0101 \in L$, so (iii) is incorrect.

1.7. Answer: (d).

- The language $A = \{ a^n b^n c^n \mid n \geq 0 \}$ is non-context-free (slide 2-96) and infinite, so (a) is incorrect. In fact, if A is non-context-free language, A must be infinite. To see why, if A were finite, then A would be regular (slide 1-95), which would imply A is context-free (Corollary 2.32).
- $A = \{ a^n b^n c^n \mid n \geq 0 \}$ is also nonregular (Corollary 2.32), so (b) is incorrect.
- $A = \{ a^n b^n c^n \mid n \geq 0 \}$ is non-context-free, and $abc \in A$ but $(abc)^{\mathcal{R}} = cba \notin A$, so A is not closed under reversals, making (c) incorrect.

1.8. Answer: (c).

- By Theorem 2.20, a language is context-free if and only if some PDA recognizes it, so we can answer the question by considering CFLs. The language $\{ a^n b^n \mid n \geq 0 \}$ is context-free but infinite, so (a) is incorrect.
- The language $\{\varepsilon\}$ is finite, so it is regular (slide 1-95), and Corollary 2.32 ensures it is also context-free, so (b) is incorrect.
- By Theorem 2.20, a language is context-free if and only if some PDA recognizes it, and Theorem 2.9 then guarantees that the language has a CFG in Chomsky normal form, so (d) is incorrect.

1.9. Answer: (c).

- Kleene's Theorem (Theorem 1.54) implies that L must be regular, so (a) is incorrect.
- Because L must be regular, Corollary 2.32 ensures L is also context-free, so (c) is correct and (b) is incorrect.
- For the language L with regular expression ab^*, we have that $x = ab \in L$ and $y = abb \in L$, but $xy = ababb \notin L$, so L is not closed under concatenation, making (d) incorrect.

1.10. Answer: (d).

- Consider $A = \{ 0^n 1^n \mid n \geq 0 \}$, which is nonregular (slide 1-105). Let $C = \{ 0^{2n+1} 1^{2n+1} \mid n \geq 0 \}$, so C is the set of strings in A with an odd number of 0s followed by exactly the same number of 1s, and C is infinite. Now let $B = A - C = \{ 0^{2n} 1^{2n} \mid n \geq 0 \}$, so B is the set of strings in A with an even number of 0s followed by exactly the same number of 1s. We can show that B is nonregular by the pumping lemma, as follows. Suppose that B is regular, and consider $s = 0^{2p} 1^{2p} \in B$, where p is the pumping length. Note that $|s| = 4p \geq p$, so the conclusions of the pumping lemma must hold. Splitting the string $s = xyz$ as in the pumping lemma leads to $x = 0^j$ for some $j \geq 0$, $y = 0^k$ for some $k \geq 1$, and $z = 0^m 0^p 1^{2p}$ for some $m \geq 0$, where $j + k + m = p$. But the pumped string $xyyz = 0^j 0^k 0^k 0^m 0^p 1^{2p} = 0^{2p+k} 1^{2p} \notin B$, which is a contradiction. Thus, B is nonregular, showing (a) is incorrect.
- Consider $A = \{ 0^n 1^n \mid n \geq 0 \}$, which is nonregular (slide 1-105), and let $C = A$, which is infinite. Then $B = A - C = \emptyset$, which is regular (B has regular expression $\emptyset$, so B is regular by Kleene's theorem), so (b) is incorrect. Also, B then is also context-free (Corollary 2.32), so (c) is incorrect.

2. (a) $(a \cup b)((a \cup b)(a \cup b))^*$.

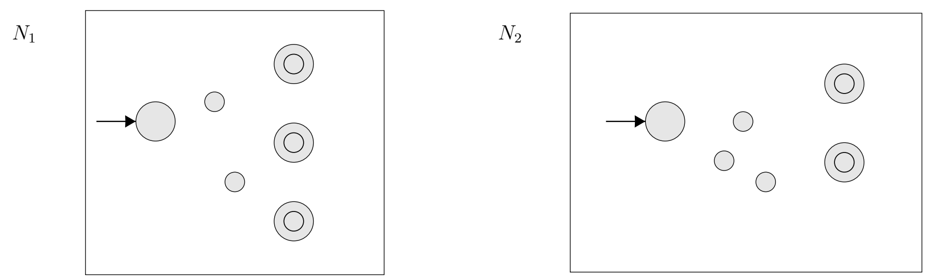
There are infinitely many other correct regular expressions for this language, such as $((a \cup b)(a \cup b))^*(a \cup b)$
or $(a \cup b)(aa \cup ab \cup ba \cup bb)^*$
or $a((a \cup b)(a \cup b))^* \cup b((a \cup b)(a \cup b))^*$ or ....
Some incorrect answers include

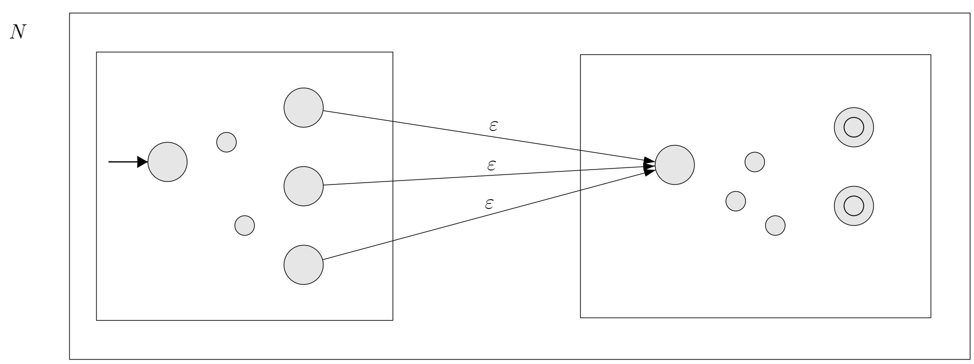
- $(a \cup b)^*((a \cup b)(a \cup b))^*$, which generates $\varepsilon \notin A$ and $ab \notin A$;

- $(a \cup b)^*(aa \cup bb)^*$, which can't generate $aba \in A$;
- $a(aa)^* \cup b(bb)^*$, which can't generate $aba \in A$;
- $((a \cup b)(a \cup b)(a \cup b))^*$, which generates $bbbbbb \notin A$;
- $(a \cup b)^n$ for n odd, which is not a regular expression.

(b) $a^*(b(a \cup b) \cup \varepsilon)b^*(a \cup \varepsilon)a^*$. Another regular expression is $a^*b(a \cup b)b^*(a \cup \varepsilon)a^* \cup a^*b^*(a \cup \varepsilon)a^*$. There are infinitely many correct regular expressions for this language.

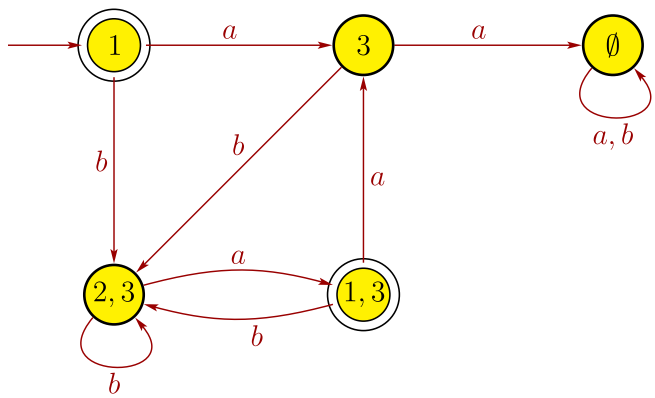
(c) As on slide 1-63 of the notes, if A_1 is defined by NFA N_1 and A_2 is defined by NFA N_2, then an NFA N for $A_3 = A_1 \circ A_2$ is as below:





(d) (Homework 5, problem 3c.) Assume that $S_3 \notin V_1$. Then a CFG for $A_3 = A_1^*$ is $G_3 = (V_3, \Sigma, R_3, S_3)$ with $V_3 = V_1 \cup \{S_3\}$ and $R_3 = R_1 \cup \{ S_3 \to S_1S_3 \mid \varepsilon \}$.

3. A DFA for C is below:



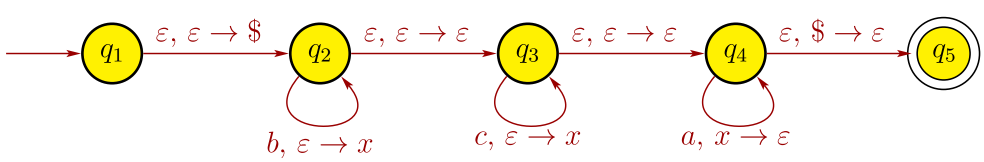
4. (a) For $\Sigma = \{a, bc\}$, let $L = \{\, b^i c^j a^k \mid i, j, k \geq 0, \text{ and } i + j = k \,\}$ be the language given in the problem. A CFG $G = (V, \Sigma, R, S)$ for L has $V = \{S, X\}$ with S the starting variable, $\Sigma = \{a, b, c\}$, and rules

$$S \to bSa \mid X$$
$$X \to cXa \mid \varepsilon$$

There are infinitely many other correct CFGs for L.

An incorrect CFG for L has rules $S \to bSa \mid cSa \mid \varepsilon$, which can derive $S \Rightarrow cSa \Rightarrow cbSaa \Rightarrow cbaa \notin L$.

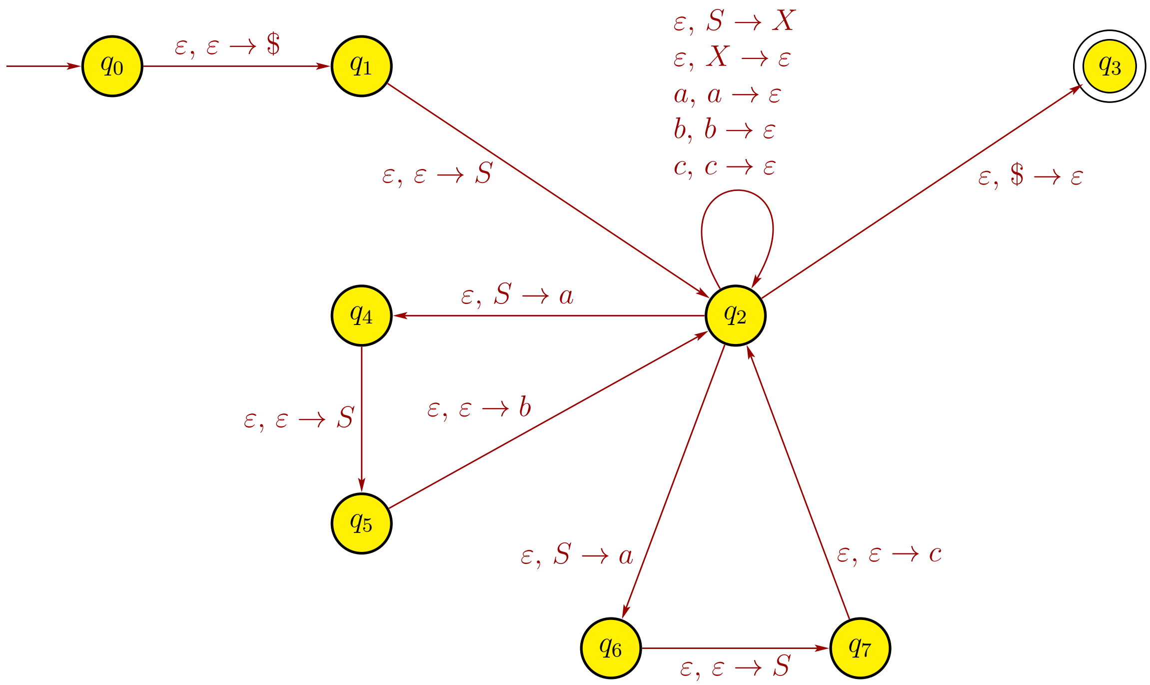
(b) There are infinitely many correct PDAs for L. Here is one:



In the above PDA,

- state q_2 pushes an x for each b read,
- state q_3 pushes an x for each c read,
- state q_4 pops an x for each a read to match the b's read in state q_2 and the c's read in state q_3,
- transition from q_4 to q_5 pops \$ to make sure stack is empty.

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



Note that

- The path $q_2 \to q_4 \to q_5 \to q_2$ corresponds to the rule $S \to bSa$, where the symbols on the right side of the rule are pushed in reverse order.
- The path $q_2 \to q_6 \to q_7 \to q_2$ corresponds to the rule $X \to cSa$, where the symbols on the right side of the rule are pushed in reverse order.

5. Language $A = \{ b^i c^j a^k \mid i, j, k \geq 0 \text{ and } i + j = k \}$ 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 $s = b^p a^p$, where $s \in A$ because $s = b^i c^j a^k$ for $i = k = p$ and $j = 0$. Also, we have that $|s| = 2p \geq p$, so the Pumping Lemma will hold. Thus, there exist 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) $|xy| \leq p$.

   Because 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 first set of b's (possibly none), 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

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

Because

$$b^p a^p = s = xyz = b^j b^k b^m a^p = b^{j+k+m} a^p,$$

we must have that

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

The first property implies that the pumped string $xy^2z \in A$, but

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

since $k \geq 1$, so in the pumped string, the sum of the number of b's and the number of c's does not equal the number of a's. This contradicts the first property of the pumping lemma. Therefore, A is a nonregular language.

Another possible string that will result in a contradiction is $s = b^p c^p a^{2p} \in A$, where $|s| = 4p > p$. Then splitting $s = xyz$ satisfying properties (ii) and (iii) of the pumping lemma will lead to

$$\begin{aligned} x &= b^j \text{ for some } j \geq 0, \\ y &= b^k \text{ for some } k \geq 1, \\ z &= b^m c^p a^{2p} \text{ for some } m \geq 0, \end{aligned}$$

where $j + k + m = p$. Property (i) of the pumping lemma states that $xyyz \in A$, but $xyyz = b^{p+k} c^p a^{2p} \notin A$ because $p + k + p \neq 2p$ since $k > 0$, giving a contradiction.