

CS 341-005 and HM1, Fall 2025, Face-to-Face Section
Solutions for Midterm 1

1. Multiple choice.

1.1. Answer: (e).

- For $\Sigma = \{a, b, c\}$, let $B = \Sigma^*$, which is regular. Now any language A having alphabet Σ satisfies $A \subseteq B$. Taking $A = \{a^n b^n c^n \mid n \geq 0\}$, which is non-context-free (slide 2-96) so it is also non-regular (Corollary 2.32), which shows that (a) and (c) are incorrect. On the other hand, let $A = \emptyset$, which is finite so A is regular (slide 1-95) and also context-free (Corollary 2.32), showing that (b) and (d) are incorrect.

1.2. 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.

1.3. Answer: (a).

- By Corollary 2.32, every regular language (even if it is infinite) is also context-free, so (a) is correct, and (b) and (c) are incorrect.
- Theorem 1.25 implies that $L_1 \cup L_2$ must be regular, so (d) is incorrect.

1.4. Answer: (e).

- For $\Sigma = \{a, b, c\}$, the language $L_1 = \{a^n b^n c^n \mid n \geq 0\}$ is infinite, not regular, and not context-free, so (a), (b), and (c) are incorrect.
- For the same language L_1 , note that $a \in \Sigma^*$ but $a \notin L_1$, making (d) incorrect.

1.5. Answer: (e).

- The regular expression $b^*(aa)^*$ generates the string $bbaa \notin A$, so (a) is incorrect.
- The regular expression $(baa)^*$ generates the string $baabaa \notin A$, so (b) is incorrect.
- The given CFG G in part (c) has language $L(G) = \emptyset$ (i.e., no strings at all) because derivations can never terminate: $S \Rightarrow bSaa \Rightarrow bbSaaaa \Rightarrow bbbSaaaaaa \Rightarrow \dots$, so (c) is incorrect.
- The language A has CFG with rules $S \rightarrow bSaa \mid \varepsilon$, so (d) is incorrect.

1.6. Answer: (c).

- The regular expression 1^* defines an infinite language, so (a) is incorrect.
- The regular expression $0 \cup 1$ defines a finite language, so (b) is incorrect.
- By Kleene's Theorem (Theorem 1.54), A must be regular, so (c) is correct.

1.7. Answer: (b).

1.8. Answer: (b).

- HW 4, problem 5c, shows that (a) is incorrect, and that (b) is correct.
- Slightly modifying the proof on slide 1-105 shows that the language $L_1 = \{a^n b^n \mid n \geq 1\}$ is non-regular. Adding ε to L_1 leads to the language $L_2 = \{a^n b^n \mid n \geq 0\}$, which is context-free (with CFG having rules $S \rightarrow aSb \mid \varepsilon$), so (c) is incorrect.
- Slightly modifying the proof on slide 2-96 shows that the language $L_1 = \{a^n b^n c^n \mid n \geq 1\}$ is non-context-free, so it is also non-regular by Corollary 2.32. Adding ε to L_1 leads to the language $L_2 = \{a^n b^n c^n \mid n \geq 0\}$, which is non-context-free by the proof on slide 2-96, so (d) is incorrect.

1.9. Answer: (b).

- The regular expression 01^*0 cannot generate the string $11 \in L$, so R1 is incorrect.
- The regular expression R2 is correct; see Example 7 on slide 1-74.
- The regular expression $(0 \cup 1)(0 \cup 1)^*(0 \cup 1)$ cannot generate the string $1 \in L$, so R3 is incorrect.

1.10. Answer: (c).

- The regular expression $(0^*1^*)^*(1^*0^*)^*$ generates $01 \notin A$, so (a) is incorrect. In fact, the language A is not regular, so A does not have a regular expression.
- The CFG can yield $S \Rightarrow 1$, but $1 \notin A$ because the string has odd length, so (b) is incorrect.
- The language A is even-length palindromes of 0's and 1's, and the CFG in part (c) is correct, as seen in HW 5, problem 1(b).

1.11. Answer: (c).

- For $\Sigma = \{0, 1\}$, consider the finite languages $A = \{\epsilon\}$ and $B = \{1\}$. Then $\overline{A} = \Sigma^* - \{\epsilon\} = \Sigma^+$ is infinite, and $\overline{A} \circ B = \Sigma^+ \circ \{1\}$, which is the set of all strings of length ≥ 2 that end in 1. Thus, $\overline{A} \circ B$ is infinite, so (a) is incorrect.
- For $\Sigma = \{0, 1\}$, consider the finite languages $A = \{\epsilon\}$ and $B = \emptyset$, so $\overline{A} = \Sigma^+$ but $\overline{A} \circ B = \emptyset$, which is finite, so (b) is incorrect.
- Because A and B are finite, both are regular (slide 1-95), and $\overline{B}$ is regular because the class of regular languages is closed under complements (slide 1-22). Also, the class of regular languages is closed under intersection, so $A \cap \overline{B}$ is regular, implying that $A \cap \overline{B}$ is also context-free (Corollary 2.32), making (c) correct. Also, $\overline{A}$ is regular, so $\overline{A} \cup \overline{B}$ is regular because the class of regular languages is closed under union (Theorem 1.25), making (d) incorrect.

1.12. Answer: (a).

- $|A| = 101$, so it is finite, making it regular (slide 1-95). Thus, A has a regular expression by Kleene's Theorem (Theorem 1.54), making (a) correct.
- Because A is regular, it must have a regular expression, so (b) is incorrect. Also, (d) is incorrect by Corollary 1.40. In addition, every regular language is also context-free (Corollary 2.32), so A must have a PDA (Theorem 2.20), making (c) incorrect.

1.13. Answer: (a).

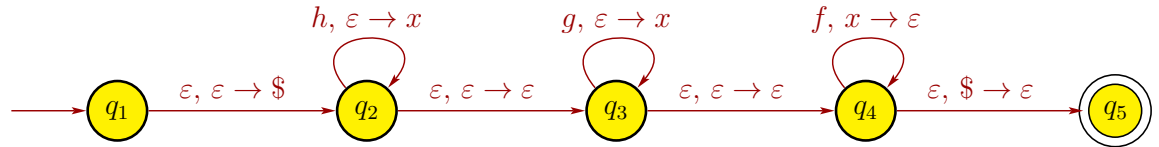
- If A is non-context-free, it must also be nonregular by Corollary 2.32, so option (a) is correct.
- The language $B = \Sigma^*$ is both context-free and regular, so option (b) is incorrect. Also, $\overline{B} = \emptyset$, which is finite, so option (c) is incorrect.
- For the alphabet $\Sigma = \{a, b, c\}$, consider the language $A = \{a^n b^n c^n \mid n \geq 0\} \subseteq \Sigma^*$, and slide 2-96 proves A is non-context-free. Define $B = \Sigma^*$, which is regular so it is also context-free by Corollary 2.32. But $A \cup B = B$ is context-free, showing that option (d) is incorrect.

1.14. Answer: (c).

- $\varepsilon \notin \emptyset$, so (a) is incorrect.
- $\emptyset$ is the empty set, and ε is the empty string, so they aren't equal, making (b) incorrect.
- $\emptyset^* = \{\varepsilon\}$ making (c) correct, and (d) incorrect.

1.15. Answer: (h).

- For $\Sigma = \{f, g, h\}$, the PDA



recognizes the language $A = \{h^i g^j f^k \mid i, j, k \geq 0 \text{ and } i + j = k\}$, which is a slight variation of HW 6, problem 1(e). Note that A differs from L_1 , L_2 , and L_3 . For example, the string $h^1 g^2 f^3 \in A$ but $h^1 g^2 f^3 \notin L_m$ for $m = 1, 2, 3$. Thus, (h) is correct.

- The languages L_1 and L_2 are context-free, so they have PDAs (HW 6, problem 1(g) is a slight variation of L_2) but just not the one that is given. On the other hand, L_3 is not context-free (slight variation of slide 2-96), so L_3 does not have a PDA (Theorem 2.20).

2. Short answers.

- (a) Let $\Sigma = \{c, d\}$ (**note the alphabet!**), and let A be the set of strings $w \in \Sigma^*$ such that $|w|$ is even and w ends in c , where $|w|$ denotes the length of w .

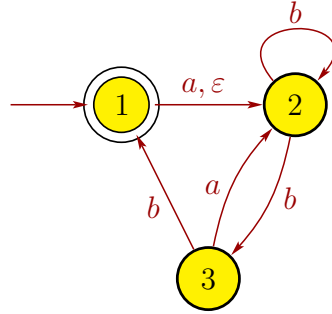
Answer: There are infinitely many correct regular expressions for A , including $((c \cup d)(c \cup d))^*(c \cup d)c$, $(\Sigma\Sigma)^*\Sigma c$, $(cc \cup cd \cup dc \cup dd)^*(cc \cup dc)$.

- (b) Suppose that language A_1 has CFG $G_1 = (V_1, \Sigma, R_1, S_1)$ and language A_2 has CFG $G_2 = (V_2, \Sigma, R_2, S_2)$. Give a CFG G_3 for A_2^* in terms of G_1 and G_2 . You do not have to prove the correctness of your CFG G_3 , but do not give just an example.

Answer: $G_3 = (V_3, \Sigma, R_3, S_3)$, where

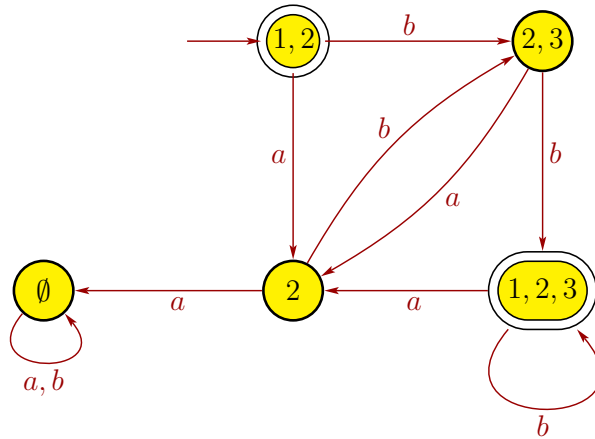
- $V_3 = V_2 \cup \{S_3\}$,
- Σ is the same as in G_2 ,
- rules $R_3 = R_2 \cup \{S_3 \rightarrow S_2 S_3, S_3 \rightarrow \varepsilon\}$,
- S_3 is the start variable, where $S_3 \notin V_2$.

3. Let N be the following NFA with $\Sigma = \{a, b\}$, and let $C = L(N)$.



Give a DFA for C . You only need to draw the state diagram (graph); do not give the 5-tuple.

Answer: A DFA for C is below (there are other correct DFAs):



4. For $\Sigma = \{a, b, c\}$, let

$$L = \{b^{2n}c^{3k}a^n \mid n \geq 0, k \geq 0\}.$$

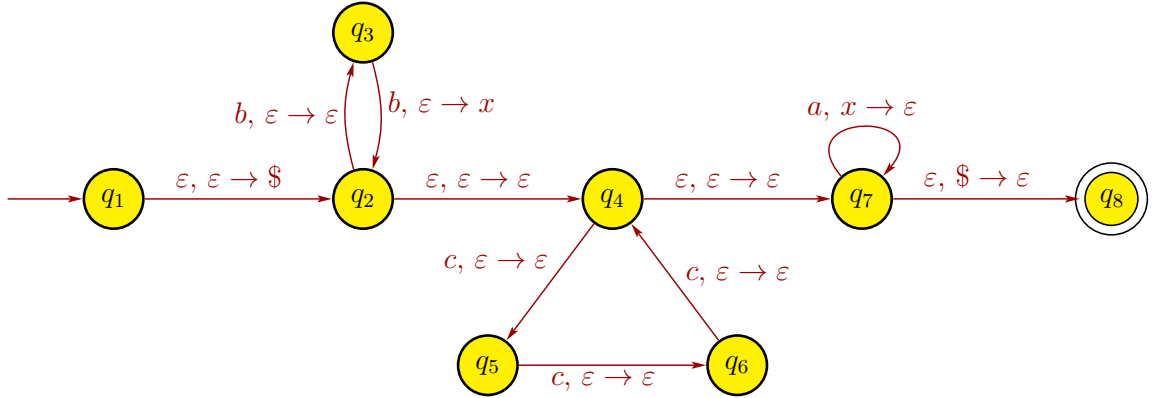
Answer: This is a slight variation of HW 6, problem 4(c).

- (a) One CFG $G = (V, \Sigma, R, S)$ for language L has a set of variables $V = \{S, T\}$, where S is the start variable; set of terminals $\Sigma = \{a, b, c\}$; and rules

$$\begin{aligned} S &\rightarrow bbSa \mid T \\ T &\rightarrow cccT \mid \varepsilon \end{aligned}$$

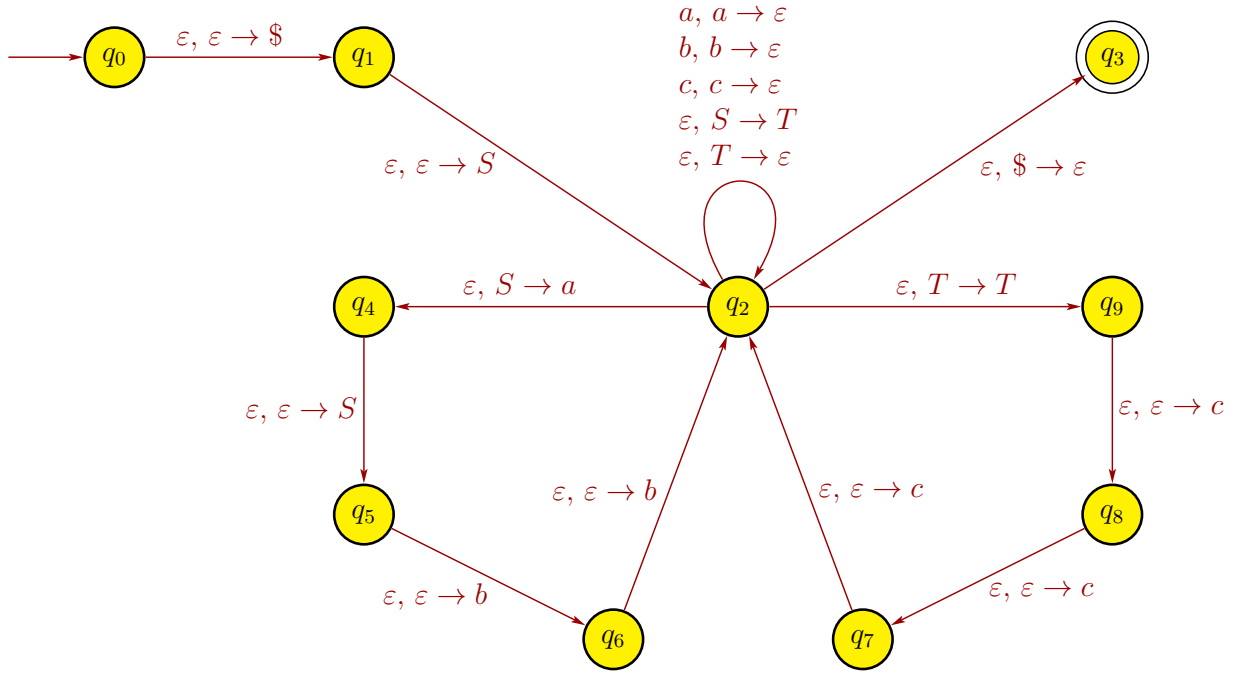
There are infinitely many other correct CFGs for L .

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



where an edge label “ $x, y \rightarrow z$ ” means read x , pop y , and push z . The state q_1 can also be an accept state.

We can also design a PDA for L by applying the algorithm from Lemma 2.21 to convert the CFG in part (a) into a PDA.



Note that

- The path $q_2 \rightarrow q_4 \rightarrow q_5 \rightarrow q_6 \rightarrow q_2$ corresponds to the rule $S \rightarrow bbSa$, where the symbols on the right side of the rule are pushed in reverse order.
- The path $q_2 \rightarrow q_9 \rightarrow q_8 \rightarrow q_7 \rightarrow q_2$ corresponds to the rule $T \rightarrow cccT$, where the symbols on the right side of the rule are pushed in reverse order.

5. For $\Sigma = \{e, f\}$ (**note the alphabet!**), let $A = \{www \mid w \in \Sigma^*\}$.

Is A a regular or nonregular language? If A is regular, give a regular expression **and** DFA (only state diagram) for A . If A is not regular, prove that it is a nonregular language.

Answer: The language A is nonregular. Note that A is from problem 3a from HW 4 but with a different Σ . To prove this, suppose that A is a regular language. Let p be the “pumping length” of the pumping lemma (Theorem 1.70). Consider the string $s = e^p f e^p f e^p f$. Note that $s \in A$ since $s = www$ with $w = e^p f$, and $|s| = 3p + 3 \geq p$, so the pumping lemma will hold. Thus, we can split the string s into 3 parts $s = xyz$ satisfying the properties

- $xy^iz \in A$ for each $i \geq 0$,
- $|y| > 0$,
- $|xy| \leq p$.

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

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

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

$$\begin{aligned} xy^2z &= e^j e^k e^k e^m fe^p fe^p f \\ &= e^{p+k} fe^p fe^p f \end{aligned}$$

since $j + k + m = p$. Hence, $xy^2z \notin A$ because it is impossible to split xy^2z into equal thirds, so we get a contradiction. Therefore, A is a nonregular language.