1. For each of the following languages, determine whether it is (I) regular, (II) not regular but context-free, or (III) not context-free. Prove your answer.

- $L_1 = \{a^ib^jc^i | i < j < k\}$.
- $L_2 = \{w \neq w^R | w \in \{a, b\}^*\}$
- $L_3 = \{a^ib^jc^i | i < j \text{ or } k < j\}$.
- $L_4 = \{a^ib^jc^i | i < j \text{ and } k < j\}$.
- $L_5 = \{a^ib^jc^i | i < j \Rightarrow k < j\}$.
- $L_6 = \{uvw | u, v, w \in \{0, 1\}^* \text{ and } |u| = |v| = |w|\}$
- $L_7 = \{a^ib^jc^i | i, j \geq 0\}$.
- $L_8 = \{a^ib^jc^i | i, j \geq 0\}$.
- $L_9 = \{a^ib^jc^i | i, j \geq 0\}$.
- $L_{10} = \{w \in \{a, b\}^* | \mu_0(w) = \mu_k(w) \text{ and } w \text{ does not contain substring } abaa \text{ or } babb\}$.
- $L_{11} = \{a^ib^jc^i | i = 0 \text{ or } j = k = 1\}$.
- $L_{12} = \{a^n|n \geq 0\} \cdot \{b^n|n \geq 0\} \cdot \{e^n|n \geq 0\}$.
- $L_{13} = \{w \in \{0, 1\}^* \text{ every substring 010 in } w \text{ is followed immediately by substring } 111\}$.
- $L_{14} = \{a^ib^j | 2i \neq 3j\}$.
- $L_{15} = \{a^ib^jc^i | i \neq j \text{ or } j \neq k \text{ or } i \neq k\}$.
- $L_{16} = \{xyz | x, y, z \in \{0, 1\}^* \text{, } |x| = |z| \text{ and } \mu_0(x) \geq \mu_0(z)\}$.
- $L_{17} = \{xyz | x, y, z \in \{0, 1\}^* \text{, } |x| = |z| > 0, \text{ and } \mu_0(x) \geq \mu_0(z)\}$.
- $L_{18} = \{0^p1^q0^m1^n | p + q = m + n, p, q, m, n \geq 0\}$.
- $L_{19} = \{x^ny^n | x, y \in \{0, 1\}^* \text{, } x \text{ is a substring of } y\}$.
- $L_{20} = \{x^ny^n | x, y \in \{0, 1\}^+\}$.
- $L_{21} = \{0^i1^j | i = 2^j\}$.
- $L_{22} = \{1^n0^n | n \geq 0\}$.
- $L_{23} = \{1^n1^n | n \geq 0\}$.
- $L_{24} = \{1^n1^n | n \geq 0\}$.
- $L_{25} = \{xy^zy^Rx | x, y, z \in \{a, b\}^*\}$.
- $L_{26} = \{a^ib^j | i \text{ and } j \text{ are two prime numbers smaller than } 1000\}$.
- $L_{27} = \{b_i b_i \}_{i \geq 1} | b_i \text{ is the binary representation of integer } i, i \geq 0\}$.
- $L_{28} = \{a^ib^j | 2i \neq 3j + 1\}$.
- $L_{29} = \{xy|x, y \in \{0, 1\}^* \text{, } \mu_0(x) = \mu_1(y)\}$.
- $L_{30} = \{w \in \{a, b, c, d\}^* | \mu_0(w) \geq \mu_c(w) \geq \mu_d(w) \geq 0\}$.
- $L_{31} = \{a^nb^nc^nd^n | n \leq m\}$.
- $L_{32} = \{w = xyz | x \in 0^*, y \in 1^*, z \in 0^*, |x| = |z|, \text{ and } |y| = 2 \cdot |z|\}$.
- $L_{33} = \{xzcz | x, z \in \{a, b\}^* \text{ and } |x| = |z|\}$.
- $L_{34} = \{a^ib^jc^i | i, j, k \geq 0 \text{ and } (i = 2j \text{ or } i = 3k)\}$.
- $L_{35} = \{w \in \{a\}^* \exists y \in \{a\}^* \text{ where } w = yy^R\}$.
- $L_{36} = \{0^n1^n0^n1^n | n \geq 0\}$.
- $L_{37} = \{0^n3^n0^n|n|n \geq 0\}$.

---

\*Feel free to address any questions about the problems to me at nakhleh@cs.utexas.edu.
1. Let $L$ be a language, and define $Op(L)$ to be
\[
\{x \in L \mid \text{no proper prefix of } x \text{ is in } L\}.
\]

(a) Describe $Op(L)$ for the following languages:

i. $L_a = \{a^n b^n | n \geq 1\}$.

ii. $L_b = \{a^n b^m | 0 \leq n, 0 \leq m, \text{ and } n + m \text{ is even}\}$.

iii. $L_c = \{a^i b^j c^k | i \leq k \text{ or } j \leq k\}$.

(b) If $L$ is regular, then $Op(L)$ is regular. Prove.

(c) If $L$ is context free, then $Op(L)$ is not necessarily context free. Prove.

2. Let $L$ be a language, and define $Op(L)$ to be
\[
\{x \in L \mid \text{no proper prefix of } x \text{ is in } L\}.
\]

(a) Describe $Op(L)$ for the following languages:

i. $L_a = \{a^n b^n | n \geq 1\}$.

ii. $L_b = \{a^n b^m | 0 \leq n, 0 \leq m, \text{ and } n + m \text{ is even}\}$.

iii. $L_c = \{a^i b^j c^k | i \leq k \text{ or } j \leq k\}$.

(b) If $L$ is regular, then $Op(L)$ is regular. Prove.

(c) If $L$ is context free, then $Op(L)$ is not necessarily context free. Prove.

3. Prove that the class of context-free languages is closed under (I) union, (II) concatenation, and (III) Kleene star.

(don’t copy&paste from your class notes; the purpose of this question is to see whether you can prove those properties on your own).

4. (a) Let $C$ be a CFL and $R$ be a regular language. Prove that the language $C \cap R$ is context free.

(b) Use part (a) to show that the language $A = \{w | w \in \{a, b, c\}^* \text{ and contains equal number of } a \text{‘s, } b \text{‘s, and } c \text{‘s} \}$ is not a CFL.

5. (a) Use the languages $A = \{a^n b^m c^n | m, n \geq 0\}$ and $B = \{a^n b^m c^n | m, n \geq 0\}$ to show that the class of CFL’s is not closed under intersection.

(b) Use part (a) and DeMorgan’s Law to show that the class of CFL’s is not closed under complementation.

6. Show that the CFL’s are closed under the following operation
\[
\text{init}(L) = \{w | \exists x, wx \in L\}
\]

7. Answer each of the following questions, and prove your answer.

(a) Let $L_1$ and $L_2$ be two CFL’s. Is it necessarily true that $L_1 - L_2$ is also a CFL?

(b) Let $L_1$ and $L_2$ be two CFL’s. Is it necessarily true that, if $L_1 = L_2 L_3$ then $L_3$ is a CFL?

(c) If $L_1$ is a CFL and $L_2 \subseteq L_1$, is it necessarily true that $L_2$ is a CFL?

(d) If $L_1$ is a CFL and $L_1 \subseteq L_2$, can $L_2$ be a regular language?

8. Convert each of the following CFG’s into Chomsky Normal Form.

\[
G_1: \quad S \rightarrow ABC \\
A \rightarrow aC|D \\
B \rightarrow bB|\varepsilon|A \\
C \rightarrow Ac|\varepsilon|Cc \\
D \rightarrow aA \\
\]

\[
G_2: \quad S \rightarrow AB \\
A \rightarrow aA|aC \\
B \rightarrow bB|bC \\
C \rightarrow \rho W \\
W \rightarrow TV \\
T \rightarrow t|\varepsilon
\]
9. Show that, if \( G \) is a CFG in Chomsky normal form, then for any string \( w \in L(G) \) of length \( n \geq 1 \), exactly \( 2n - 1 \) steps are required for any derivation of \( w \).

10. Consider the following CFG whose only variable is \( E \):

\[
E \rightarrow E \land E \lor E \lor \neg(E) \lor E \lor \neg(E) \lor p \lor q
\]

(a) Show two different parse trees for the string \( p \land q \lor p \).
(b) Show a PDA that accepts \( L(G) \).

11. (a) Show that the following CFG, \( G \), is ambiguous:

\[
E \rightarrow E + E \lor (E \lor E) \lor E \lor (E \div E) \lor (E) \lor a \lor b
\]

(b) Find an equivalent unambiguous grammar for \( L(G) \).

12. Show that \( L = \{a^i b^j c^k \mid i \neq j \text{ or } i \neq k \} \) is not a DCFL.

13. Show that \( L = \{xy \mid x, y \in \{a, b\}^* \text{ and } x \neq y \} \) is a CFL.

14. Show that \( L = \{xy \mid x, y \in \{0, 1\}^* \text{ and } |x| = |y| \text{ but } x \neq y \} \) is a CFL.

15. (I) Given a CFG \( G \) and a string \( w \), give an algorithm (procedure) to decide whether:

(a) \( L(G) \) is empty.
(b) \( L(G) \) is finite.
(c) \( L(G) \) is infinite.
(d) \( w \in L(G) \).

(II) Given a regular expression \( E \) and a PDA \( M \), give an algorithm to decide whether the language accepted by \( M \) is a subset of the language generated by \( E \).

16. Show that the following languages are DCFL’s:

(a) \( L_A = \{w w^R \mid w \in a^* b^* \} \).
(b) \( L_B = \{w \mid w \in \{a, b\}^* \text{ and each prefix of } w \text{ has at least as many } a \text{'s as } b \text{'s} \} \).
(c) \( L_C = \{w \mid w \in \{a, b\}^* \text{ and } m \cdot \#_a(w) = n \cdot \#_b(w) \text{ for arbitrary constants } m \text{ and } n \} \)

17. For each of the following languages, describe (the state diagram) of a PDA that recognizes the language:
(a) \(\{0^n w_1^m | w \in \{0,1\}^*, \#_0(w) = \#_1(w) \text{ and } m \geq n\}\).
(b) \(\{0^n w_1^m | w \in \{2,3\}^*, \#_2(w) = \#_3(w) \text{ and } m \geq n\}\).
(c) \(\{w \in \{0,1\}^* | \forall x \in \{0,1\}^*, \text{ if } w = xy \text{ for some } y \in \{0,1\}^*, \text{ then } \#_0(x) \geq \#_1(x)\}\).
(d) \(\{w1^n | w \in \{a,b\}^*, \text{ and } \#_a(w) = n\}\).

18. More challenging problems:

(a) Prove that the following language is CF:

\[ A = \{w \in \{a,b\}^* | w \neq xx \text{ for any } x \in \{0,1\}^*\} \]

(b) Construct a PDA that accepts the language

\[ B = \{w \in \{0,1\}^* | 2\#_0(w) \neq 3\#_1(w)\} \]