

“It is in middles that extremes clash, where ambiguity restlessly rules.” – John Updike

1. Given below is a list of production rules of context free grammars. Identify the language accepted by them.

$$(a) \begin{aligned} S &\rightarrow aSa \mid aBa \\ B &\rightarrow bB \mid b \end{aligned}$$

$$(b) \begin{aligned} S &\rightarrow TU \\ T &\rightarrow 0T1 \mid \epsilon \\ U &\rightarrow 1U0 \mid \epsilon \end{aligned}$$

$$(c) \begin{aligned} A &\rightarrow BAB \mid B \mid \epsilon \\ B &\rightarrow 00 \mid \epsilon \end{aligned}$$

2. Come up with context free grammars for the following languages:

$$(a) \{w \in \{0, 1\}^* \mid |w|_1 = 2|w|_0\}$$

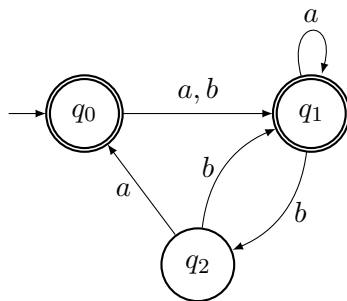
$$(b) \{0^i 1^j 2^k \mid i \neq j \text{ or } j \neq k\}$$

$$(c) \{w\#x \mid w^R \text{ is a substring of } x, \text{ where } w, x \in \{a, b\}^*\}$$

$$(d) \{a, b\}^* - \text{palindromes}$$

**Bonus** .  $\{t_1\#t_2\#\dots\#t_k \mid \forall i, t_i \in \{a, b\}^*, t_i \neq t_j, \text{ for some } i \neq j\}$

3. Formally prove the correctness of your answers for 1(b) and 2(a).  
4. (a) Find a context-free grammar that generates the language accepted by the following finite automaton:



- (b) Given any finite automaton  $A$ , can you give a context free grammar  $G$  such that  $\mathcal{L}(G) = \mathcal{L}(A)$ ?
5. A *right linear grammar* is a context-free grammar in which every production has at most one non-terminal on the right hand side, and this non-terminal appears at the right end of the production; i.e., every production rule is of the form  $A \rightarrow a_1 \dots a_k B$  or  $A \rightarrow a_1 \dots a_k$ . Show that right linear grammars generate precisely the class of regular languages.
6. Identify which of the following grammars are ambiguous. Can you come up with an unambiguous grammar for them? If not, try to argue why not.

$$(a) A \rightarrow A + A \mid A - A \mid a$$

- (b) Statement  $\rightarrow$  **if** Condition **then** Statement **else** Statement | **if** Condition **then** Statement | atomic  
Condition  $\rightarrow$  atomic

This is the typical evaluation of if-then-else statements in programming. *atomic* denotes statements that are simply executed, and do not include if-then-else within them. If this notation is uncomfortable, just consider the grammar with one non-terminal  $S$ , and  $\Sigma = \{i, t, e, a\}$ , defined as:

$$S \rightarrow iatSeS \mid iatS \mid a$$

- (c) The grammar you constructed for 2(b).