1. Let $A$ be the language consisting of those strings $w$ in $\{0, 1, \#\}^*$ such that either $w$ starts with 0 or $w = u\#u$ for some $u \in \{0, 1\}^*$. $A$ is not regular. In a Pumping Lemma proof of this, you are given $p$ and you choose $s$. For each of the following possible choices of $s$, state whether or not the choice is a good one. If the choice is bad, provide the decomposition that allows the string to be pumped.

(a) $s = 0^p1^p\#0^p1^p$

Solution: This is bad choice when $p \geq 2$ because if we take $x = \varepsilon$, $y = 0$, $z = 0^{p−1}1^p\#0^p1^p$, then $xy^iz$ starts with a 0 for all values of $i$, so is in the language.

(b) $s = 1^p0^p\#1^p0^p$

Solution: This is a good choice. Given $x, y, z$ as in the Pumping Lemma, we choose $i = 2$. Then, $xy^iz$ starts with a 1, and the part to the left of the $\#$ is different from the part to the right, so the string is not in the language.

(c) $s = (10)^p\#(10)^p$

Solution: This is a good choice for the same reason as in Part (b).

2. Read the discussion of minimum pumping length given in Problem 1.55 of the text (third US edition) and then give the minimum pumping length for the following languages. Justify your answers.

(a) $0^*1^*0^+$

Solution: The minimum pumping length is 2. To see this, first note that $p = 1$ is not a pumping length because 0 is in the language and it cannot be pumped to stay in the language because if $0 = xyz$ with $|y| > 0$, then $y = 0$ and $xz = \varepsilon$ is not in the language.

Now suppose that $s$ is in the language and has length $\geq 2$. Then, $s$ has the form $s = 0^n1^m0^r$ with $r > 0$. If $n > 0$, we can let $x = \varepsilon$, $y = 0$ and $z = 0^{n−1}1^m0^r$. Then $|y| = 1$, $|xy| = 1 \leq 2$ and $xy^iz$ is in the language for all $i \geq 0$. If $n = 0$ and $m > 0$, then we let $x = \varepsilon$, $y = 1$ and $z = 1^{m−1}0^r$. We have $|y| > 0$, $|xy| = 1$ and $xy^iz$ is in the language for all $i \geq 0$. If $n = 0$ and $m = 0$, then, since $|s| \geq 2$, we must have $r > 1$, so we can let $x = \varepsilon$, $y = 0$, and $z = 0^{r−1}$. Then, $|y| = 1$, $|xy| = 1 \leq 2$, and $xy^iz$ is in the language for all $i \geq 0$. Thus, 2 is a pumping length for this language.

(b) $\{0^n1^m|n + m \text{ is divisible by 3}\}$

Solution: The minimum pumping length is 6. To see this, we first note that 5 is not a pumping length. This is because $s = 0010000$ is a string in the language with length 7, but if $s = xyz$ with $|y| > 0$ and $|xy| \leq 5$, then either $y$ contains 1 or $y$ consists of one or two zeros.
In the first case, pumping $y$ gives a string which has more than one 1, so is not in the language, while in the second case, $xyyz$ is a string with seven or eight 0’s, so $xyyz$ is not in the language.

To see that 6 is a pumping length, let $s$ be in the language with $|s| \geq 6$. Then, $s = 0^n10^m$ for some $n$ and $m$ with $n + m$ divisible by 3. If $n \geq 3$, we chose $x = \varepsilon, y = 000$ and $z = 0^{n-3}10^m$. We have $|y| > 0$, $|xy| = 3 \leq 6$, and $xy^iz$ is in the language for all $i \geq 0$. If $n < 3$, then, since $|s| \geq 6$, we must have $m \geq 3$, so we take $x = 0^n1$, $y = 000$, $z = 0^{m-3}$. Then, $|y| > 0$, $|xy| = n + 4 \leq 6$, and $xy^iz$ is in the language for all $i \geq 0$.

(c) $\{w \in \{0, 1\}^*| |w| \leq 5\}$

**Solution:** The minimum pumping length is 6. This is because 5 is not a pumping length since 0$^5$ is in the language and cannot be pumped to stay in the language, while there are no strings of length 6 or more in the language, so every string of length 6 or more that is in the language can be pumped.

3. **Problem 1.54**

(a) To see that $F$ is not regular, by Problem 1.31, it suffices to show that $F^R$ is not regular. We do this by showing that $F^R$ does not have the Pumping Property. Given $p$, choose $s = c^pbpa$. Then $s \in F^R$ and $|s| \geq p$. Given $x, y, z$ as in the Pumping Lemma, we must have $y = c^k$ for some $k > 0$. Choose $i = 2$. Then, $xyyz = c^{p+k}b^ka$, which is not in $F^R$.

(b) Choose $p = 2$. We show that this is a pumping length for $F$. Given $s \in F$ with $|s| \geq 2$, we must have we $s = a^ib^jc^k$. We consider cases depending on what $i, j, k$ are.

Case 1: $i = 0, j = 0$. Then we must have $k > 0$. Take $x = \varepsilon, y = c, z = c^{k-1}$. Then $|xy| \leq p$, $|y| > 0$ and for all $r \geq 0$, $xy^rz = c^{k+r-1}$, which is in $F$.

Case 2: $i = 0, j > 0$. Take $x = \varepsilon, y = b, z = b^{j-1}c^k$. Then $|xy| \leq p$, $|y| > 0$ and for all $r \geq 0$, $xy^rz = b^{j+r-1}c^k$, which is in $F$.

Case 3: $i = 1$. Then we must have $j = k$. Take $x = \varepsilon, y = a, z = bjc$. Then $|xy| \leq p$, $|y| > 0$ and for all $r \geq 0$, $xy^rz = a^{i+r-1}b^jc^k$, which is in $F$.

Case 4: $i = 2$. Take $x = \varepsilon, y = aa, z = bjc^k$. Then $|xy| \leq p$, $|y| > 0$ and for all $r \geq 0$, $xy^rz = a^{2r}b^jc^k$, which is in $F$ (since $2r \neq 1$).

Case 5: $i > 2$. Take $x = \varepsilon, y = a, z = a^{i-1}bjc^k$. Then $|xy| \leq p$, $|y| > 0$ and for all $r \geq 0$, $xy^rz = a^{i+r-1}b^jc^k$, which is in $F$ (since $i + r - 1 \neq 1$).

(Note that the special treatment in Case 4 is necessary because if $s = aabjc^k$ we can’t take $y = a$ since then $xz = abjc^k$ is not in $F$ if $j \neq k$.)
(c) The Pumping Lemma states that every regular language has the Pumping Property. It does not state that every language with the Pumping Property is regular, so the result in this problem does not contradict the Pumping Lemma.

4. Give context-free grammars for the following languages:
   (a) \{0^n1^n|n \geq 0\}.
   \[
   S \rightarrow 0S11|\varepsilon
   \]
   (b) \{0^n1^m|n < m\}.
   \[
   S \rightarrow 0S1T \\
   T \rightarrow 1T1
   \]
   (c) \{w\#u|w, u \in \{0, 1\}^* \text{ and } |w| > 2|u|\}.
   \[
   S \rightarrow BBSB|BT\# \\
   T \rightarrow BT|\varepsilon \\
   B \rightarrow 0|1
   \]
   (d) \{x_1\#x_2\#x_3|x_1, x_2, x_3 \in \{a, b\}^* \text{ and either } x_3 = x_1^R \text{ or } x_3 = x_2^R\}.
   \[
   S \rightarrow V\#T|U \\
   T \rightarrow aTa|bTb|\# \\
   U \rightarrow aUa|bUb|\#V\# \\
   V \rightarrow aV|bV|\varepsilon
   \]

5. In class, we gave the following two grammars for the language
   \[
   L = \{w|w \text{ contains the same number of } a's \text{ and } b's\}.
   \]
   \[
   S \rightarrow \varepsilon|aB|bA \\
   A \rightarrow aS|bAA \\
   B \rightarrow bS|aBB
   \]
   and
   \[
   S \rightarrow aSb|bSa|SS|\varepsilon
   \]
   For each of these grammars, give both a leftmost derivation and a parse tree for the string \textit{aababbab}.

   **Solution:**
   A leftmost derivation in the first grammar is:
A leftmost derivation on the second grammar is:

\[ S \Rightarrow aSb \Rightarrow aSSb \Rightarrow aabSb \Rightarrow aabSSb \Rightarrow aabaSb \Rightarrow aababSb \Rightarrow aababab \]

Parse trees corresponding to these derivations are:
6. Show that the first grammar given in the previous problem (the one with three variables) is ambiguous.

**Solution:**
The string $aababb$ has the following two leftmost derivations, so the grammar is ambiguous.

$$S \Rightarrow aB \Rightarrow aaBB \Rightarrow aabSB \Rightarrow aabB \Rightarrow aabaBB \Rightarrow aababSB \Rightarrow aababS \Rightarrow aababb$$

and

$$S \Rightarrow aB \Rightarrow aaBB \Rightarrow aabSB \Rightarrow aabaBB \Rightarrow aababSB \Rightarrow aababB \Rightarrow aababbS \Rightarrow aababb$$

7. (a) Give a context-free grammar that generates the language $A = \{ w \in \{a, b\}^* | w \text{ has more } a\text{'s than } b\text{'s} \}$

[Hint: In class, we gave two grammars that generate the set of strings with the same number of $a$’s as $b$’s. You can use either one of these grammars as part of the grammar you give.]

**Solution:**

$$
S \rightarrow TaS|TaT \\
T \rightarrow aT|bT|\epsilon
$$
(b) Give a justification that your grammar is correct. Use the same type of argument as we gave in class that the grammars that generate strings with the same number of $a$’s and $b$’s are correct. You do not need to reprove the correctness of these grammars.

**Solution:** We know from class that the variable $T$ derives exactly the strings of terminals with the same number of $a$’s as $b$’s. Suppose now that a string $w$ of terminals has more $a$’s than $b$’s. We need to analyze the string and see that it matches one of the two rules in the grammar with $S$ on the left side.

We consider a counter that counts each $a$ as +1 and each $b$ as −1. At the end of the string $w$, the count is at least 1, so we can divide up $w$ as $ucv$ where $u, v \in \{a, b\}^*$ and $c \in \{a, b\}$ and the count first reaches +1 after $c$ is read. If $c = b$, then the count would be 2 at the end of $u$, so the count would have to be +1 somewhere in the middle of $u$, contradicting the assumption that the count first reaches +1 when $c$ is read. Thus, $c = a$, so $w = uav$, and the count is 0 at the end of $u$. This means that $u$ has the same number of $a$’s as $b$’s, and hence it can be derived from $T$. We now consider two cases based on what the count is at the end of $w$.

**Case 1:** The count is +1 at the end of $w$. Then $v$ takes the count from +1 to +1, meaning that $v$ also contains the same number of $a$’s as $b$’s. Since $w = uav$, the rule $S \rightarrow TaT$ covers this case.

**Case 2:** The count is more than +1 at the end of $w$. Then, $v$ takes the count from +1 to a number greater than +1. This means that $v$ has more $a$’s than $b$’s and $w = uav$, so the rule $S \rightarrow TaS$ covers this case.

(A full proof would also require showing by induction on the length of the derivation that if a string of terminals $w$ is derivable from $S$, then $w$ has more $a$’s than $b$’s, but I was not asking for this.)

8. Give an unambiguous grammar for the language $L$ of Problem 5.

(This is a difficult problem, but give it a try. As a hint, you can use two variables other than the start symbol. One variable generates strings with the same number of $a$’s as $b$’s that have the additional property that every prefix has at least as many $a$’s as $b$’s, and the second variable generates all strings with the same number of $a$’s as $b$’s that have the additional property that every prefix has at least as many $b$’s as $a$’s.)

**Solution:**

$$
S \rightarrow \varepsilon | aTbS | bUaS \\
T \rightarrow \varepsilon | aTbT \\
U \rightarrow \varepsilon | bUaU
$$