1. Consider integers written in base 3 with no leading 0s. Let $L_1$ be the set of such strings which represent odd numbers.

(a) Construct a DFA that accepts $L_1$.

answer:

(b) Construct a left-linear grammar for $L_1$.

answer: Let $Y, B, R$ denote the yellow, blue, and red states resp. Associate with each state the set of strings whose computation ends at that state. Then the language is composed of those strings associated with $Y$. The derivation rules of the grammar are

\[
\begin{align*}
B & \rightarrow \lambda \\
Y & \rightarrow B1 + R1 + Y(0 + 2) \\
R & \rightarrow B2 + Y1 + R(0 + 2)
\end{align*}
\]

The initial non-terminal of the grammar is $Y$. 


2. Consider the language $L_2$ generated by the following grammar
\[
S \rightarrow AB + C \\
A \rightarrow aB + C \\
B \rightarrow Ab + C \\
C \rightarrow b + aaaC
\]

Characterize $L_1$ using a combination of set notation and regular expressions.

**Answer**: (This problem was harder than intended because of a typo - mea culpa).

- $C$ generates the regular language $(aaa)^*b$; From now on we use $C$ to denote this language.
- Substituting the rules for $B$ into the rules for $A$ we obtain
  \[
  A \rightarrow a(\lambda + A) + C = aAb + (a + \lambda)C.
  \]
  Thus, $A$ generates the language
  \[
  \{a^n(a + \lambda)Cb^n | n \geq 0\}.
  \]
- Similarly, $B$ generates the language
  \[
  \{a^nC(b + \lambda)b^n | n \geq 0\}.
  \]
- Finally, $A$ generates the concatenation of $A$ and $B$, union the language $C$:
  \[
  \{a^n(a + \lambda)Cb^n | n \geq 0\}\{a^nC(b + \lambda)b^n | n \geq 0\} + C
  \]
  where $C = (aaa)^*b$.

3. What does it mean for an infinite set to be “countable”?

**Answer**: An infinite set $S$ is countable if and only there exists a bijection between $S$ and the natural numbers.
4. Construct a DFA equivalent to the following NFA.
5. Consider the language over $\Sigma = \{a, b, c\}$ consisting of strings with more occurrences of the pattern “abc” than occurrences of the pattern “abb”. Is this a regular language? Justify your answer.

**answer:** It is not a regular language. This can be proven using the Pumping Lemma:

- the word $x = (abc)^{n+1}(abb)^n$ is in the language for any $n$;
- for large enough $n$, $x = uvw$ such that
  - $uv$ is a prefix of $(abc)^{n+1}$ with $v$ not equal to $\lambda$;
  - $uv^i w$ is in the language for all $i \geq 0$;
- in particular, the word $uw^0 w = uw$ is in the language;
- but this erases at least one symbol from a prefix of $(abc)^{n+1}$;
- the resulting word does not contain more substrings of the form $abc$ than $abb$ and hence is not in the language, contradiction.