1. Let $L$ denote the language over \{0, 1\} which is recognized by the following NFA.

\[\begin{array}{ccc}
I & 0 \\
\varepsilon & A & 1 \\
\varepsilon & B & 0
\end{array}\]

(i) Construct a DFA which accepts $L$.
(ii) Describe the Myhill–Nerode equivalence classes for the language $L$.
(iii) Write down a grammar which defines $L$.
(iv) Find a closed form for the generating function $\sum_{n=0}^{\infty} \ell_n z^n$, where $\ell_n$ is the number of strings of length $n$ in $L$.

**Solution**

(i) The $\varepsilon$-closure of $I$ is $T = \{I, A, B\}$ and this becomes the start state for the DFA. The other states of the DFA are the $\varepsilon$-closures of the non-empty subsets of states of the NFA. In fact we only need $\{A\}$, $\{B\}$ and $\emptyset$. Thus the DFA is

\[\begin{array}{ccc}
I & 0 \\
\varepsilon & A & 1 \\
\varepsilon & B & 0
\end{array}\]

We can see directly that this DFA accepts $L$ by noting that $L$ is defined by the regular expression $0^* + 1^*$. That is the strings in $L$ have all 0's or all 1's.

(ii) Recall that strings $s_1$ and $s_2$ are equivalent whenever for all strings $u$ we have $s_1u \in L$ if and only if $s_2u \in L$. The empty string is in $L$ and it forms an equivalence class by itself. The collection of strings of one or more 0's is a single equivalence class as the collection of strings of one or more 1's. The remaining strings have a mixture of 0's and 1's and form a single equivalence class (corresponding to a dead-end state). Thus there are four equivalence classes.

(iii) A right linear grammar $G$ which generates $L$ exists as $L$ is accepted by a DFA. For convenience, label the states $T$, $A$, $B$ and $\emptyset$, of the DFA, as $S$, $X$, $Y$ and $Z$.
respectively. In $G$ take the set of non-terminal symbols to be \{S, X, Y, Z\}, with S as the start symbol. The set of terminal symbols is \{0, 1\} and the productions are:

\[
\begin{align*}
S &\rightarrow \epsilon | 0X | 1Y \\
X &\rightarrow \epsilon | 0X | 1Z \\
Y &\rightarrow \epsilon | 1Y | 0Z \\
Z &\rightarrow 0Z | 1Z
\end{align*}
\]

\(iv\) We have $\ell_n = 2$ except that $\ell_0 = 1$ and therefore the generating function is $2(1 - z)^{-1} - 1$.

2. Let $L$ be the language over the alphabet \{0, 1\} consisting of all strings which end in 111.

\(i\) Design an NFA which recognizes $L$.

\(ii\) Apply the subset construction to produce a DFA which recognizes $L$.

**Solution**

\(i\) A possible solution is

\[
\begin{array}{c}
\text{0, 1} \\
\text{I} & \\
1 & 1 & 1 & \\
A & B & C
\end{array}
\]

\(ii\) The subset construction produces the DFA

\[
\begin{array}{c}
\text{D} & \\
0 & 1 & 0 & 1 \\
I & 1 & 0 & 0 & 1 & 1 & \\
E & F
\end{array}
\]

where $D = \{I, A\}$, $E = \{I, A, B\}$, $F = \{I, A, B, C\}$.

3. Describe the languages generated by the following grammars. The start symbol is $S$.

\(i\) The non-terminal symbols are $S, A$ and $B$, the terminal symbols are 0 and 1 and the productions are

\[
\begin{align*}
S &\rightarrow ABS \\
A &\rightarrow 0 \\
B &\rightarrow 1 \\
S &\rightarrow \epsilon.
\end{align*}
\]

\(ii\) The non-terminal symbol is $S$, the terminal symbol is $a$ and the productions are

\[
\begin{align*}
S &\rightarrow aaaS \\
S &\rightarrow \epsilon.
\end{align*}
\]
**Solution**

In both cases the language is regular and can be described by a regular expression. In (i) the regular expression is \((01)^*\) and in (ii) it is \((aaa)^*\).

4. **Give a grammar that has the given set of strings as its language.**

   (i) All strings over \(\Sigma = \{p\}\) consisting of an even number of \(p\)'s
   
   (ii) All strings over \(\Sigma = \{p, q, r\}\) consisting of \(n\) \(p\)'s followed by \(n\) \(q\)'s, followed by a single \(r\), where \(n = 1, 2, 3, \ldots\).

**Solution**

(i) Non-terminal symbol \(S\), terminal symbol \(p\) and productions \(S \rightarrow ppS\) and \(S \rightarrow \varepsilon\).

(ii) Non-terminal symbols \(S\) and \(D\), terminal symbols \(p, q\) and \(r\) and productions

\[
\begin{align*}
S & \rightarrow Dr \\
D & \rightarrow pDq \\
D & \rightarrow pq
\end{align*}
\]