MH4300		**Name:**	Singh Jasraj
Assignment 2		**Matric:**	U1940558D
11 October 2022		**Collaborators:**	Tan Hui Ru Kimberley

Problem 1

$$\begin{aligned}
Q^{(a,b)}(z) &= \sum_{t \in \mathcal{T}_{(a,b)}} \mathrm{xpl}(t) z^{|t|} \\
&= \sum_{t_L \in \mathcal{T}_{(a,b)}} \sum_{t_R \in \mathcal{T}_{(a,b)}} \left[\mathrm{xpl}(t_L) + \mathrm{xpl}(t_R) + a\left(|t_L|+1\right) + b\left(|t_R|+1\right)\right] z^{|t_L|+|t_R|+1} \\
&= z \sum_{t_L \in \mathcal{T}_{(a,b)}} \sum_{t_R \in \mathcal{T}_{(a,b)}} \left[\mathrm{xpl}(t_L) + \mathrm{xpl}(t_R) + a|t_L| + b|t_R| + a + b\right] z^{|t_L|} z^{|t_R|} \\
&= 2zT(z)Q^{(a,b)}(z) + (a+b)z^2 T(z)T'(z) + (a+b)zT(z)^2
\end{aligned}$$

where

$$T(z) = \frac{1-\sqrt{1-4z}}{2z}$$

is the generating function for the Catalan numbers. Now, we can solve for $Q(z)$.

$$\begin{aligned}
Q^{(a,b)}(z) &= (a+b)zT(z)\frac{zT'(z)+T(z)}{1-2zT(z)} \\
&= \frac{a+b}{4}\left[\frac{2zT(z)(2zT(z))'}{1-2zT(z)}\right] \\
&= \frac{a+b}{4}\left[\frac{1-\sqrt{1-4z}}{\sqrt{1-4z}}\right]\left[\frac{2}{\sqrt{1-4z}}\right] \\
&= \frac{a+b}{2}\left[\frac{1-\sqrt{1-4z}}{1-4z}\right]
\end{aligned}$$

Problem 2

$$\begin{aligned}
F_0 &= 1 \\
F_1 &= 2 \\
F_2 &= F_0 + F_1 \\
F_3 &= F_1 + F_2 \\
&\vdots \\
F_N &= F_{N-1} + F_{N-2} \\
&\vdots \\
\Rightarrow F(z) &= 1 + 2z + z^2 F(z) + z\left(F(z) - 1\right) \\
&= 1 + z + (z + z^2)F(z) \\
\Rightarrow F(z) &= \frac{1+z}{1-z-z^2}
\end{aligned}$$

We observe that the sum of path lengths to external nodes for a tree is the sum for its left sub-tree, $\mathrm{xpl}(f_{N-1})$, and right sub-tree, $\mathrm{xpl}(f_{N-2})$, and the additional length coming from the traversal of the edge joining the root node to the two sub-trees, $1 \times \mathrm{x}(f_{N-1}) + 2 \times \mathrm{x}(f_{N-2})$, where $\mathrm{x}(\cdot)$ gives the number of external nodes.

$$\begin{aligned}
\mathrm{xpl}(f_N) &= \mathrm{xpl}(f_{N-1}) + \mathrm{xpl}(f_{N-2}) + 1 \times \mathrm{x}(f_{N-1}) + 2 \times \mathrm{x}(f_{N-2}) \\
G_N &= G_{N-1} + G_{N-2} + F_{N-1} + 2F_{N-2}
\end{aligned}$$

Now, we solve this recursion.

$$\begin{aligned}
G_0 &= 0 \\
G_1 &= 3 \\
G_2 &= G_1 + G_0 + F_1 + 2F_0 \\
G_3 &= G_2 + G_1 + F_2 + 2F_1 \\
&\vdots \\
G_N &= G_{N-1} + G_{N-2} + F_{N-1} + 2F_{N-2} \\
&\vdots \\
\Rightarrow G(z) &= 3z + zG(z) + z^2 G(z) + z\left(F(z) - 1\right) + 2z^2 F(z) \\
&= \frac{2z + zF(z) + 2z^2 F(z)}{1 - z - z^2} \\
&= \frac{z(3+z)}{(1-z-z^2)^2}
\end{aligned}$$

Problem 3

(a) To design a valid mapping for the Oompa Loompas, we need to ensure that the encoding for no symbol is the prefix of encoding for another symbol. This can be achieved using binary trees. Specifically, we can use a binary tree with 55 external nodes, one for each symbol, where each edge connecting a node to its left child is labelled by a dot, ·, and edge connecting a node to its right child labelled by a dash, -. We map each external node to a symbol, and the encoding for a symbol is defined as the concatenation of the dots and dashes on the path connecting the root to the corresponding external node.

We can prove that this is a valid mapping using contradiction- if the encoding for one of the symbols was a prefix for another symbol, we would have that the former symbol is on the path to the latter, which is not possible since the symbols are mapped to external nodes of the binary tree.

(b) Assuming that the transmission time is only dependent on the length of the message being transmitted, comparing the two encoding strategies involves comparing the expected length of encoding for a random symbol. Since we are using binary trees and assume that each symbol is equally likely to appear in the message, the problem reduces to comparing xpl$(\cdot)$ for the trees corresponding to the two strategies.

Let xpl(A, N) denote the average external path length of strategy A and xpl(B, N) denote the external path length of strategy B, for tree with N external nodes (or $N-1$ internal nodes).

$$\begin{aligned}\text{xpl}(A, 55) &= \frac{[z^{54}]Q^{(1,2)}(z)}{[z^{54}]T(z)} \\ &\approx \frac{4.494911 \times 10^{32}}{4.519597 \times 10^{29}} \\ &\approx 994.538008\end{aligned}$$

where,

$$\begin{aligned}Q^{(a,b)}(z) &= \frac{a+b}{2}\left[\frac{1-\sqrt{1-4z}}{1-4z}\right] \\ T(z) &= \frac{1-\sqrt{1-4z}}{2z}\end{aligned}$$

Noting that $[z^8]F(z) = 55$,

$$\begin{aligned}\text{xpl}(B, 55) &= [z^8]G(z) \\ &= 461\end{aligned}$$

Since xpl$(A, 55) >$ xpl$(B, 55)$, strategy B is better for transmission.