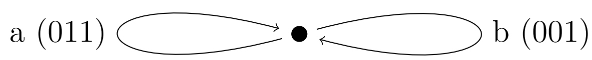
Homework 2

Information Theory for Complex Systems

a)

We can construct an optimal code using two symbols, $a = 011$ and $b = 001$ transforming the automata into an equivalent one presented below.



The compressed automata generates a random sequence of "coin flips" between a and b. Since it is completely uniformly random, binary sequence the entropy per symbol is $s_{a,b} = 1\,\text{bit}$. The original sequence of 1's and 0's contains three symbol per a or b so the resulting entropy per symbol is $s = 1/3\,\text{bit}$.

b)

Every sequence longer than 3 necessarily contains at least one copy of a or b allowing one to identify where in the automata the current state is located. Therefore $k_m = 0,\ \forall\, m > 3$.

c)

Equation (3.36) gives

$$\eta = \sum_{m=1}^{\infty} (m-1) k_m. \tag{1}$$

Following equation (3.18) gives

$$k_1 = p(0)\log(2p(0)) + p(1)\log(2p(1)) = \frac{1}{2}\log(1) + \frac{1}{2}\log(1) = 0. \tag{2}$$

We can then

$$\begin{aligned} k_2 = p(0)\left[p(0|0)\log\frac{p(0|0)}{p(0)} + p(1|0)\log\frac{p(1|0)}{p(0)}\right] + \\ p(1)\left[p(0|1)\log\frac{p(0|1)}{p(1)} + p(1|0)\log\frac{p(1|1)}{p(1)}\right] = \\ \frac{1}{2}\left[\frac{1}{3}\log\frac{1/3}{1/2} + \frac{2}{3}\log\frac{2/3}{1/2}\right] + \left[\frac{2}{3}\log\frac{4}{3} + \frac{1}{3}\log\frac{2}{3}\right] \approx 0.0817... \end{aligned} \tag{3}$$

We also have the equation

$$k_{corr} = S_{max} - s = 1 - \frac{1}{3} \iff k_{corr} = \frac{2}{3} = k_1 + k_2 + k_3 \iff k_3 \approx 0.58496 \tag{4}$$

Finally reinserting for η gives $\eta \approx 1.25162$.