Homework 2

Information Theory for Complex Systems

Vincent Udén

February 2023

1 New and better idea

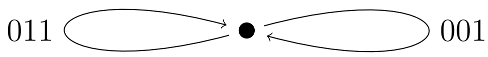

Figure 1: Optimal code representation of the finitite state automata.

2 Old idea

A labeled finite automata generates a sequences of 0s and 1s. It is not a Markov system since the probability distribution for the next symbol is not immediately known from the current symbol. Knowing the current symbol being 0 can result in 100% chance of generating a 1 and 100% chance of generating a 0, depending on where in the system one is located.

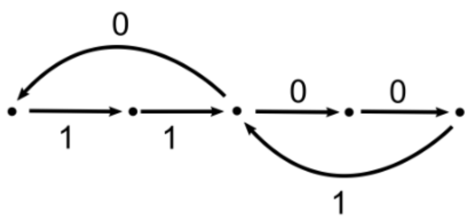

Figure 2: The finite automata in question.

We introduce names to the 5 nodes, in order from left to right: A, B, C, D and E. The entropy per symbol is calculated using

$$s = \lim_{m\to\infty} \sum_{x_1 \ldots x_{m-1}} p(x_1 \ldots x_{m-1}) \sum_{x_m} p(x_m|x_1 \ldots x_{m-1}) \log \frac{1}{p(x_1 \ldots x_{m-1})}, \quad (1)$$

from the book. For convenince's sake, we can introduce a variable $z = x_1 \ldots x_{m-1}$, resulting in

$$s = \lim_{m\to\infty} \sum_{z} p(z) \sum_{x_m} p(x_m|z) \log \frac{1}{p(z)}. \quad (2)$$

Next up, we need to calculate the probability of being at each node in the system. As defined by the excercise, when situated at a node with several exiting edges, one is picked at random with equal probability. Therefore, the probability of ending up at a node is proportional to the amount of edges pointing to it. For normalisation, all probabilities must of course sum to 1. This results in the following system of equations

$$\begin{cases} p(A) = p(B) = p(D) = p(E) \\ p(C) = 2p(A) \\ p(A) + p(B) + p(C) + p(D) + p(E) = 1. \end{cases} \quad (3)$$

Substituting the two upper equations into the bottom gives

$$p(C) + p(C) + p(C) = 3p(C) = 1 \iff p(C) = \frac{1}{3}, \quad (4)$$

which in turn gives

$$p(A) = p(B) = p(D) = p(E) = \frac{1}{6}. \quad (5)$$

The entropy we're looking for doesn't consider the made up nodes A, B, C, D, E but the 0s and 1s which they generate. Therefore we also need to establish the probability of getting a 0 or 1 from every transition between the nodes. Most of the nodes only have one exiting edge and the last one has two exiting nodes with equal probability. Therefore

$$\begin{array}{ll} p(1|A) = 1 & p(0|A) = 0 \\ p(1|B) = 1 & p(0|B) = 0 \\ p(1|C) = 0 & p(0|C) = 1 \\ p(1|D) = 0 & p(0|D) = 1 \\ p(1|E) = 1 & p(0|E) = 0 \end{array} \quad (6)$$

all nodes generate a desterministic symbol, even though the edge chosen is random. Therefore equation (2) simplifies to

$$s = \lim_{m \to \infty} \sum_{z} p(z) \log \frac{1}{p(z)}, \tag{7}$$

Is it possible to convert this to a system of 2 symbols? 011 and 001. This would turn the system into a Markov system, reducing the complexity of the calculation. The entropy would only depend on

$$s = p(B) \left[p + ++ \right] \tag{8}$$