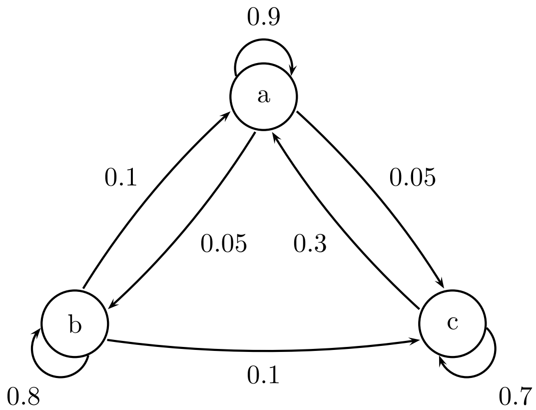
Markov source, example

Markov source

Suppose we have a stationary Markov source X_n of order 1 with the alphabet $\mathcal{A} = \{a, b, c\}$ and the transition probabilities $P(x_n|x_{n-1})$ given by

$$\begin{array}{lll} P(a|a) = 0.9 & P(b|a) = 0.05 & P(c|a) = 0.05 \\ P(a|b) = 0.1 & P(b|b) = 0.8 & P(c|b) = 0.1 \\ P(a|c) = 0.3 & P(b|c) = 0 & P(c|c) = 0.7 \end{array}$$

We can draw the Markov source as the state graph below:



Stationary distribution

The stationary distribution of the Markov source are the probabilities $P(x_n)$ for being in the states at any given time. For a source of order 1 this is the same as the probabilities for the different symbols. For the example source this can be gotten from the equation system

$$\begin{cases} P(a) &= 0.9 \cdot P(a) + 0.1 \cdot P(b) + 0.3 \cdot P(c) \\ P(b) &= 0.05 \cdot P(a) + 0.8 \cdot P(b) \\ P(c) &= 0.05 \cdot P(a) + 0.1 \cdot P(b) + 0.7 \cdot P(c) \end{cases}$$

This is an underdetermined system with a parametric solution. By adding the equation

$$P(a) + P(b) + P(c) = 1$$

we can find the solution

$$P(a) = \frac{4}{6} \ ; \ P(b) = \frac{1}{6} \ ; \ P(c) = \frac{1}{6}$$

Entropies

The calculated probabilities immediately give us the entropies for single symbols

$$H(X_n) = -\frac{4}{6} \cdot \log \frac{4}{6} - \frac{1}{6} \cdot \log \frac{1}{6} - \frac{1}{6} \cdot \log \frac{1}{6} \approx \underline{1.2516}$$

What we are really interested in is the *entropy rate* of the source. For an order 1 Markov source like this, the entropy rate is given by $H(X_n|X_{n-1})$. To calculate this entropy, we first need to cacluate the probability distribution $P(x_{n-1}, x_n)$ for pairs of symbols from the source. For this we use Bayes' rule $P(x_{n-1}, x_n) = P(x_{n-1}) \cdot P(x_n|x_{n-1})$ which gives us

$$\begin{array}{lll} P(a,a) = 36/60 & P(a,b) = 2/60 & P(a,c) = 2/60 \\ P(b,a) = 1/60 & P(b,b) = 8/60 & P(b,c) = 1/60 \\ P(c,a) = 3/60 & P(c,b) = 0 & P(c,c) = 7/60 \end{array}$$

From this distribution we can caclulate the joint entropy

$$\begin{aligned} H(X_{n-1}, X_n) &= -36/60 \cdot \log 36/60 - 2/60 \cdot \log 2/60 - 2/60 \cdot \log 2/60 \\ &\quad -1/60 \cdot \log 1/60 - 8/60 \cdot \log 8/60 - 1/60 \cdot \log 1/60 \\ &\quad -3/60 \cdot \log 3/60 - 0 \cdot \log 0 - 7/60 \cdot \log 7/60 \\ &\approx \underline{1.9315} \end{aligned}$$

which in turn will gives us the entropy rate, using the chain rule backwards

$$H(X_n|X_{n-1}) = H(X_{n-1}, X_n) - H(X_{n-1}) \approx \underline{0.6799}$$

Note that since the source is stationary, $H(X_n) = H(X_{n-1})$. We can also calculate the entropy rate directly from the definition of conditional entropy

$$\begin{aligned} H(X_n|X_{n-1}) &= -\sum P(x_{n-1}, x_n) \cdot \log P(x_n|x_{n-1}) \\ &= -36/60 \cdot \log 0.9 - 2/60 \cdot \log 0.05 - 2/60 \cdot \log 0.05 \\ &\quad -1/60 \cdot \log 0.1 - 8/60 \cdot \log 0.8 - 1/60 \cdot \log 0.1 \\ &\quad -3/60 \cdot \log 0.3 - 0 \cdot \log 0 - 7/60 \cdot \log 0.7 \\ &\approx \underline{0.6799} \end{aligned}$$

A third way to calculate the entropy rate is a weighted average of the entropies of the outgoing probabilities for each state

$$\begin{aligned} H(X_n|X_{n-1}) &= \frac{4}{6} \cdot (-0.9 \cdot \log 0.9 - 0.05 \cdot \log 0.05 - 0.05 \cdot \log 0.05) \\ &\quad + \frac{1}{6} \cdot (-0.1 \cdot \log 0.1 - 0.8 \cdot \log 0.8 - 0.1 \cdot \log 0.1) \\ &\quad + \frac{1}{6} \cdot (-0.3 \cdot \log 0.3 - 0 \cdot \log 0 - 0.7 \cdot \log 0.7) \\ &\approx \underline{0.6799} \end{aligned}$$

The entropy rate of the source gives a lower bound on the rate when doing lossless coding of the output from the source.