COMP6237 – Tutorial

Information Theory Problem Sets

Shoaib Ehsan

s.ehsan@soton.ac.uk

Today’s Tutorial

- Problem 1
- Problem 2

Problem 1

Consider two independent integer-valued random variables, X and Y . Variable X takes on only the values of the eight integers 1, 2, ..., 8 and does so with uniform probability. Variable Y may take the value of any positive integer k, with probabilities $P_Y(k) = 2^{-k}$, k = 1, 2, 3, ...

(a) Which random variable has greater uncertainty? Calculate both entropies H(X) and H(Y).

(b) What is the joint entropy H(X,Y) of these random variables, and what is their mutual information I(X;Y)?

Solution (a)

The entropy formula is:

$$H(X) = -\sum_{i=1}^{8} p_i \log_2 p_i$$

Since it's **uniform**, $p_i = \frac{1}{8}$ for all $i = 1, 2, ..., 8$.

So,

$$H(X) = -\sum_{i=1}^{8} \frac{1}{8} \log_2 \left(\frac{1}{8}\right) = -8 \cdot \frac{1}{8} \cdot \log_2 \left(\frac{1}{8}\right) = -\log_2 \left(\frac{1}{8}\right) = \log_2(8) = 3 \text{ bits}$$

Solution (a)

$$H(Y) = -\sum_{k=1}^{\infty} P_Y(k) \log_2 P_Y(k) = -\sum_{k=1}^{\infty} 2^{-k} \cdot \log_2(2^{-k}) = \sum_{k=1}^{\infty} 2^{-k} \cdot k$$

We can use a known identity:

$$\sum_{k=1}^{\infty} kx^k = \frac{x}{(1-x)^2}, \quad \text{for } |x| < 1$$

Set $x = \frac{1}{2}$:

$$\sum_{k=1}^{\infty} k \left(\frac{1}{2}\right)^k = \frac{\frac{1}{2}}{(1-\frac{1}{2})^2} = \frac{1/2}{(1/2)^2} = \frac{1/2}{1/4} = 2$$

Solution (a)

This confirms that the entropy of Y, given by

$$H(Y) = \sum_{k=1}^{\infty} 2^{-k} \cdot k = 2 \text{ bits}$$

- H(X) = 3 bits
- H(Y) = 2 bits

→ X has greater uncertainty.

Solution (b)

Since **X and Y are independent**:

$$H(X,Y) = H(X) + H(Y) = 3 + 2 = 5 \text{ bits}$$

Mutual information is:

$$I(X;Y) = H(X) + H(Y) - H(X,Y) = 3 + 2 - 5 = 0$$

This is expected because **X and Y are independent** → they share no information.

Problem 2

Polynesian languages are famous for their small alphabets. Assume a language with the following letters and relative frequencies: p(1/8), t(1/4), k(1/8), a(1/4), i(1/8), u(1/8). What is the per-character entropy for this language? Design an (optimal, i.e. short) code to transmit a letter.

Solution

We start by developing a set of yes/no questions to identify which letter has been sampled. For this purpose, we use the divide and conquer strategy discussed in the lecture, always halving the probability mass. This procedure is not unique, so you might make other choices than I have made here. For example, one could proceed as illustrated in Figure 1.

Solution

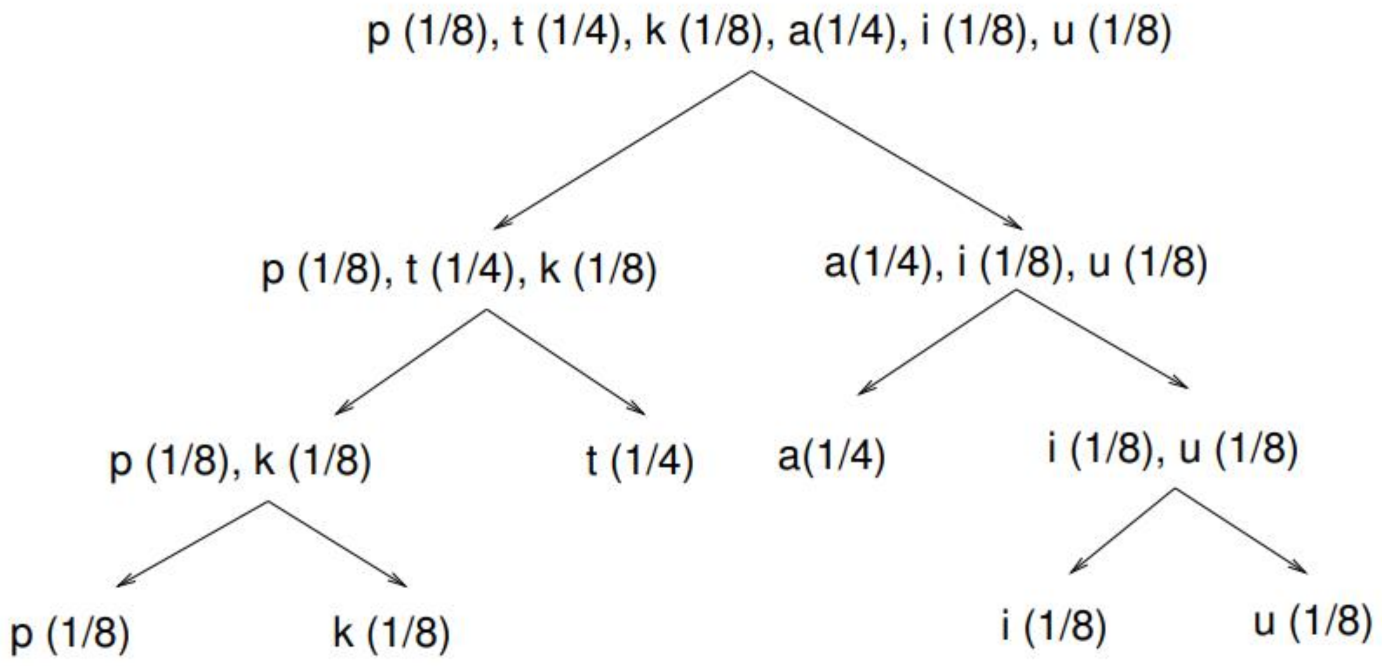

Figure 1: Figure illustrating the divide and conquer strategy applied to construct an optimal set of questions to encode the Polynesian alphabet.

Solution

Thus, the first question is: Is the letter p, t, or k? If yes, we proceed on the left-hand branch of the tree and next ask: Is it p or k? etc. The information content of each letter then corresponds to the number of questions we have to ask to identify this letter (or just $-\ln 1/p(x)$ for letter x).

This yields the entropy:

$$S = (1/4)\log 4 + (1/4)\log 4 + (1/8)\log 8 + (1/8)\log 8 + (1/8)\log 8 + (1/8)\log 8$$

$$S = 2 \times (1/4)\log 4 + 4 \times (1/8)\log 8 = 1/2(\log 4 + \log 8) = 5/2$$

Solution

Optimal codes can also be read from the tree given in Figure 1. For my choice of questions, I obtain (and depending on your questions you might have obtained a different code, but with the same code length for each symbol):

p – 111, t – 10, k – 110, a – 01, i – 001, u – 000.