COMP2610/COMP6261 – Information Theory

Tutorial 8

Zhifeng Tang (zhifeng.tang@anu.edu.au)

Question 1.

Bad codes. Which of these codes cannot be Huffman codes for any probability assignment?

(a) {0, 10, 11}

(b) {00, 01, 10, 110}

(c) {01, 10}

Question 2.

Optimal code lengths that require one bit above entropy. The source coding theorem shows that the optimal code for a random variable X has an expected length less than $H(X) + 1$. Give an example of a random variable for which the expected length of the optimal code is close to $H(X) + 1$ [i.e., for any $\epsilon > 0$, construct a distribution for which the optimal code has $L > H(X) + 1 - \epsilon$].

Question 3.

Huffman codes

(a) Construct a binary Huffman code for the following distribution on five symbols: $\mathbf{p} = (0.3, 0.3, 0.2, 0.1, 0.1)$. What is the average length of this code?

(b) Construct a probability distribution $\mathbf{p}'$ on five symbols for which the code that you constructed in part (a) has an average length (under $\mathbf{p}'$) equal to its entropy $H(\mathbf{p}')$.

Question 4.

Optimal codes. Let $l_1, l_2, \ldots, l_{10}$ be the binary Huffman codeword lengths for the probabilities $p_1 \geq p_2 \geq \ldots \geq p_{10}$. Suppose we get a new distribution by splitting the last probability mass. What can you say about the optimal binary codeword lengths $\tilde{l_1}, \tilde{l_2}, \ldots, \tilde{l_{11}}$ for the probabilities $p_1, p_2, \ldots, p_9, \alpha p_{10}, (1-\alpha)p_{10}$, where $0 \leq \alpha \leq 1$.