Information Theory
Retake of the First Midterm
October 30, 2019

1) Give the definition of uniquely decodable codes.

2) State McMillan's theorem.

3) We have a prefix code over a 7-element alphabet that has n codewords with lengths $l_1, l_2, \ldots, l_n$, respectively. Is it true then that a binary prefix code also exists with n codewords and respective codeword lengths $3l_1, 3l_2, \ldots 3l_n$? Answer the same question in the two slightly different situations, too, when the first alphabet has 8 elements or 9 elements, respectively.

4) We toss a fair coin 5 times. Let T denote the number of tails among the results of these tosses. Give an optimal average length binary encoding of T.

5) Let A and B be two independent random variables, both taking its values from the set $\{2, 3, 5, 7\}$ according to the uniform distribution. Let W denote the random variable given as $W = A \cdot B$, that is the product of A and B. What is the value of the entropies $H(W)$ and $H(W|A)$?

6) Let X, Y, Z be three random variables, each taking its values on the set $\{0, 1\}$. We know that $H(X, Y, Z) = \frac{5}{2}, H(Y) = \frac{5}{8}, H(Y|X) = \frac{1}{2}$. Determine the value of $H(X|Y)$ and $H(Z)$.

3. By McMillan's theorem the condition implies that $\sum_{i=1}^{n} 7^{-l_i} \leq 1$. This implies $\sum_{i=1}^{n} 2^{-3l_i} = \sum_{i=1}^{n} 8^{-l_i} < \sum_{i=1}^{n} 7^{-l_i} \leq 1$ and so by Kraft's theorem a binary prefix code with lengths $3l_1, \ldots .3l_n$ also exists. The above calculation shows that the same is true also if we have a prefix code over an 8-element alphabet with lengths $l_1, \ldots, l_n$. However, if the condition is for a 9-element alphabet, then we only know $\sum_{i=1}^{n} 9^{-l_i} \leq 1$ which does not imply $\sum_{i=1}^{n} 2^{-3l_i} \leq 1$, so then we may not have a binary prefix code with codeword lengths $3l_1, \ldots .3l_n$.

4. Every fixed outcome of the five tosses has probability $\frac{1}{2^5} = \frac{1}{32}$. There is one outcome with exactly 0 tails and also with exactly 5 tails, so $T = 0$ and $T = 5$ both have probability $\frac{1}{32}$. There are $5 = \binom{5}{1}$ possible outcomes with exactly 1 tail and similarly with one head, that is 4 tails. So $P(T = 1) = P(T = 4) = \frac{5}{32}$.

There are $10 = \binom{5}{2}$ possible outcomes with exactly 2 tails and similarly with two heads, that is 3 tails. So $P(T = 2) = P(T = 3) = \frac{10}{32}$.
Thus we have to make an optimal code, that is a Huffman code for the distribution $(\frac{1}{32}, \frac{1}{32}, \frac{5}{32}, \frac{5}{32}, \frac{10}{32}, \frac{10}{32})$. Using the method learnt we obtain the code with codewords: 10 (for $T = 2$), 11 (for $T = 3$), 01 (for $T = 4$), 001 (for $T = 1$), 0000 (for $T = 0$) and 0001 (for $T = 5$).

5. Since the four possible values of A and B are distinct prime numbers, we have $\binom{4}{2} + 4 = 10$ options for the value of $W = A \cdot B$ and each determines the two numbers whose product is that number. If A is known then W can take four values, each with probability $\frac{1}{4}$, so $H(W|A) = \log 4 = 2$. On the other hand, the 10 possible values of W do not all have the same probability: those that are the product of two distinct numbers can occur in two ways, according to whether A takes the value of one of the two primes and B the other or vice versa, while those that are the product of one of our primes with itself, can occur only if both A and B takes this value. Thus four possible values (these are $2^2, 3^2, 5^2, 7^2$) have probability $\frac{1}{16}$ and the other six possible values have probability $\frac{2}{16} = \frac{1}{8}$. Therefore $H(W) = 6 \cdot \frac{1}{8} \log 8 + 4 \cdot \frac{1}{16} \log 16 = \frac{9}{4} + 1 = \frac{13}{4}$.

6. By the Chain rule we have

$$H(X, Y, Z) = H(X) + H(Y|X) + H(Z|X, Y) \leq H(X) + H(Y|X) + H(Z)$$

$$= H(X) + H(Z) + \frac{1}{2} \leq \frac{5}{2}.$$

Here the first inequality follows from conditioning not increasing entropy and the second from the fact that the entropy of a binary variable cannot be more than 1. Since we know that $H(X, Y, Z) = \frac{5}{2}$, this means that both inequalities should be equalities. This already implies $H(Z) = 1$ and also $H(X) = 1$ follows. Then $H(X|Y) = H(X, Y) - H(Y) = H(X) + H(Y|X) - H(Y) = 1 + \frac{1}{2} - \frac{5}{8} = \frac{7}{8}$.