1. Let $\Omega$ be the sample space of all sequences of 4 flips of a coin which is unfair, having probabilities $P(\text{heads}) = \frac{2}{3}$, $P(\text{tails}) = \frac{1}{3}$. Let $X$ be the random variable on $\Omega$ whose value is the number of heads which appear among the 4 flips. Let $Y$ be the random variable whose value is number of heads appearing among the first 2 flips.

(a) (5 points) Compute the entropy $H(X)$ of the random variable $X$.

(b) (10 points) Compute the conditional entropy $H(X|Y)$.

2. Let $W$ be a memoryless source that emits three words $\{A, B, C\}$ with probabilities $P(A) = \frac{5}{8}$, $P(B) = \frac{1}{4}$, $P(C) = \frac{1}{8}$. Consider the second extension $W^{(2)}$ of this source.

(a) (5 points) Compute the entropy $H(W^{(2)})$ for this second extension.

(b) (10 points) Compute a binary Huffman code $\mathcal{H}$ for this second extension $W^{(2)}$.

(c) (5 points) Compute the average codeword length for $\mathcal{H}$.

3. A comma code of size $t$ for a source uses the codewords

$$\mathcal{C} = \{0, 10, 110, 1110, \ldots, \underbrace{1\cdots10}_{t-1 \text{ letters}}, \underbrace{1\cdots11}_{t-1 \text{ letters}}\}$$

and assigns these words to the source words in decreasing order of their probability. The name comes from thinking of 0 as a comma.

(a) (5 points) Prove that every comma code is uniquely decipherable.

(b) (10 points) Assuming all $t$ source words have equal probability, compute the average length of a comma code of size $t$. Your answer should be a simple function of $t$ that involves no summations, only multiplications and divisions.
4. (a) (10 points) Let $W$ be a memoryless source emitting two words \{0, 1\} with probabilities $p, 1-p$ for some $p$ in $[0, 1]$. Use calculus to show that the entropy $H(W) = H(p)$ is maximized as a function of $p$ when $p = \frac{1}{2}$. What is the maximum value of $H(p)$?

(b) (10 points) Let $n$ be a real number greater than 1. Use calculus to find the value of $p$ that maximizes the function $f(p) := -p \log_n(p) = p \log_n\left(\frac{1}{p}\right)$ for $p$ in $[0, 1]$. What is the maximum value of $f(p)$?

5. Consider sending a code that contains all binary words of length $n$ through a binary symmetric channel with error probability $p$,

- first without any added parity check bit, and
- then with an added parity check bit, making all the words have length $n + 1$ and an even number of ones.

(a) (5 points) Compute the probability of an undetected error (that is, any error at all) in the first situation, without any parity check bit, as a function of $p$ and $n$. Your final answer should involve no summations.

(b) (10 points) Explain carefully why the probability of an undetected error after adding the parity check bit is exactly

$$\sum_{k=1}^{n+1} \binom{n+1}{2k} p^{2k} (1-p)^{n+1-2k}.$$ 

6. (15 points) Prove that any binary Huffman code $H$ with codewords of lengths $(\ell_1, \ldots, \ell_t)$ will always attain equality in McMillan’s inequality, that is, it will satisfy

$$\sum_{i=1}^t \frac{1}{2^{\ell_i}} = 1.$$ 

(Possible hints:
(a) This really has little to do with the probabilities of the source.
(b) Proof by induction on $t$?)