Information Theory, Coding and Cryptography

Homework Set 2

Please send your answers – either typeset on a computer, or handwritten (but readable) and scanned, preferably in PDF format – to `bouman@cwi.nl` *before Monday March 7th, 23h59. Do not forget to put your name on the first page. Good luck!*

1 Kraft's Inequality

Below, six binary codes are shown for the source symbols $x_1, \ldots, x_4$.

	Code A	Code B	Code C	Code D	Code E	Code F
x_1	00	0	0	0	1	1
x_2	01	10	11	100	01	10
x_3	10	11	100	110	001	100
x_4	11	110	110	111	0001	1000

a) [2 points] Which codes fulfill the Kraft inequality?

b) [2 points] Is a code that satisfies this inequality always uniquely decodable?

c) [2 points] Which codes are prefix-free codes?

d) [2 points] Which codes are uniquely decodable?

2 Huffman Coding

Jane, a student, regularly sends a message to her parents via a binary channel. The binary channel is lossless (i.e. error-free), but the per-bit costs are quite high, so she wants to send as few bits as possible. Each time, she selects one message out of a finite set of possible messages and sends it over the channel. There are 7 possible messages:

1. "Everything is fine"
2. "I am short on money; please send me some"
3. "I'll come home this weekend"
4. "I am ill, please come and pick me up"
5. "My study is going well, I passed an exam (...and send me more money)"
6. "I have a new boyfriend"
7. "I have bought new shoes"

Based on counting the types of 100 of her past messages, the empirical probabilities of the different messages are:

m	1	2	3	4	5	6	7
$P_M(m)$	19/100	40/100	12/100	2/100	16/100	4/100	7/100

Jane wants to minimize the average number of bits needed to communicate to her parents (with respect to the empirical probability model above).

a) [2 points] Design a Huffman code for Jane and draw the binary tree that belongs to it.

b) [4 points] For a binary source X with $P_X(0) = \frac{1}{8}$ and $P_X(1) = \frac{7}{8}$, design a Huffman code for blocks of $N = 1, 2$ and 3 bits. For each of the four codes, compute the average codeword length and divide it by N, in order to compare it to the optimal length, i.e. the entropy of the source.

c) [1 point] If you were asked at (b) to design a Huffman code for a block of $N = 100$ bits, what problem would you run into?

Consider the random variable Z with

z	1	2	3	4	5	6
$P_Z(z)$	1/10	3/10	2/10	2/100	1/10	1/10

d) [2 points] Find a *ternary* Huffman encoding for Z (i.e., using an alphabet with three symbols).

3 Arithmetic Coding

a) [4 points] Prove Proposition 8 from the lecture notes.

b) [4 points] Show that for the improved version of the arithmetic code, AC_1, it holds that

$$\ell_{AC_1}(X_1 \cdots X_n) \leq H(X_1)+H(X_2|X_1)+H(X_3|X_2)+\cdots+H(X_n|X_{n-1})+1.$$

4 Channel Coding

Consider a binary channel that flips the input bit with probability p, i.e. for input bit X and output bit Y:

$$P_{Y|X}(0|0) = 1-p, \quad P_{Y|X}(0|1) = p,$$
$$P_{Y|X}(1|1) = 1-p, \quad P_{Y|X}(1|0) = p.$$

Let X^N be a vector of N uniformly random bits. We transmit X^N (bitwise) over this channel, yielding Y^N.

a) [2 points] Determine $H(Y^N)$ and $H(Y^N|X^N)$.

b) [1 point] Find the probability that no errors occur during this transmission, i.e. $P[Y^N = X^N]$.

c) [2 points] Let $k \in \mathbb{N}$ and $k \leq N$. Find the probability that *at most* k errors occur.

To decrease the error probability, we will use the R_3 repetition code: for every bit b, we transmit the triple bbb over the channel.

d) [2 points] Describe why it would not make sense to use a R_n code for n even. Next, describe an appropriate "decoding rule" at the receiver (how to map back from, possibly perturbed, triples of bits to single bits) such that the code can correct up to one bit error (per triple).

e) [2 points] Compute the probability of bit error in case the R_3 code is used.

5 Another Kind of Entropy

In this exercise we consider a different entropy notion. Let X and Y be random variables with joint probability distribution P_{XY}. The *guessing probability* and the *min-entropy* of X are respectively defined as

$$\text{Guess}(X) := \max_x P_X(x) \quad \text{and} \quad H_{\min}(X) := -\log \text{Guess}(X).$$

The *conditional guessing probability* and the *conditional min-entropy* of X are respectively defined as

$$\text{Guess}(X|Y) := \sum_y P_Y(y)\text{Guess}(X|Y=y)$$

and

$$H_{\min}(X|Y) := -\log \text{Guess}(X|Y).$$

a) [2 points] If X has no uncertainty (i.e. $H(X) = 0$), what is $H_{\min}(X)$?

b) [2 points] If X is uniformly distributed over $\mathcal{X}$, what is $H_{\min}(X)$?

c) [4 points] Prove that $H_{\min}(XY) \geq H_{\min}(X)$.

d) [4 points] Prove that $H_{\min}(X) \geq H_{\min}(X|Y)$.

e) [4 points] Prove that $H_{\min}(X|Y) \geq H_{\min}(XY) - \log|\mathcal{Y}|$.