Work for Information Theory Supervision II

Supervision 2 covers symbol codes (continued), error correcting codes and correlated random variables. Questions are drawn from a variety of sources: past Tripos questions, MacKay, and some of my own.

Please submit the work by 18:30 the day before the supervision.

Questions

1. Explain what is meant by an optimal binary symbol code.
2. (a) Suppose we wish to generate random numbers from a non-uniform distribution $\{p_0, p_1\} = \{0.99, 0.01\}$. Compare the following two techniques. Roughly how many random bits will each method use to generate a thousand samples from this sparse distribution?
    i. The standard method: use a standard random number generator to generate an integer between 1 and 2^{32}. Rescale the integer to $(0, 1)$. Test whether this uniformly distributed random variable is less than 0.99 and emit a 0 or a 1 accordingly.
    ii. Arithmetic coding using the correct model, fed with standard random bits.

   (b) Use adaptive arithmetic coding (Laplace model) to encode in decimal the string `DEADBEEF#` (where `#` is the end-of-string symbol). The 5-symbol source alphabet is `A`, `B`, `D`, `E`, `F`. Use a fixed probability of 0.05 for the end-of-string symbol.
3. (a) (**Optional** – if you want practice at LZ encoding.) Encode the string `00000000000010000000000` using the basic Lempel-Ziv algorithm presented in lectures. Use a 3-bit fixed-length dictionary; ignore termination of the string.

   (b) **Hard.** Using a variable-length dictionary (max. 3 bits) initialised to `0` $=$ `0`, `1` $=$ `1` and ignoring end-of-string markers, decode the LZ-encoded string: `00110010001110100010`.

   (c) Give examples of simple sources that have low entropy but would not be compressed well by the Lempel-Ziv algorithm.
4. Consider a (7, 4) Hamming code which maps $k = 4$ information bits to a length $n = 7$ codeword. Assume 0 and 1 are equiprobable in the input data.

   Use the convention used by the rest of the world, not the one presented in lectures. The transmitted codeword is

   $$[b_1, b_2, b_3, b_4, b_5, b_6, b_7]$$

   where b_3, b_5, b_6 and b_7 are the source bits and

   $$\begin{aligned} b_4 &= b_5 \oplus b_6 \oplus b_7 \\ b_2 &= b_3 \oplus b_6 \oplus b_7 \\ b_1 &= b_3 \oplus b_5 \oplus b_7. \end{aligned}$$

   (We'll discuss the reasons for using this convention in the supervision.)

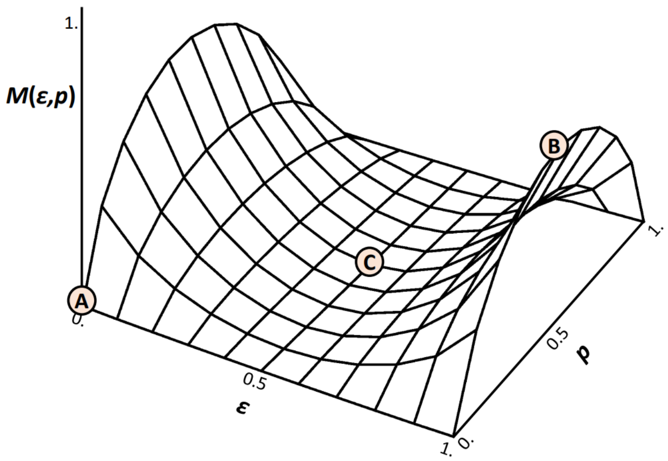
   (a) Suppose that a codeword is transmitted over a binary symmetric channel (BSC) and the received codeword is $r = [1, 1, 0, 1, 0, 1, 1]$. Decode the received sequence to a codeword.

   (b) Calculate the probability of block error p_B of the (7, 4) Hamming code as a function of the bit error p and show that to leading order it goes as $21p^2$.

(c) If the (7, 4) Hamming code can correct any one bit error, might there be a (14, 8) code that can correct any two errors?

5. Consider using the repetition code R_5 to encode binary input symbols for transmission through a binary symmetric channel with $f = 0.3$. Assuming $p_0 < 0.5$, find the maximum value of p_0 for which the optimal decoder's rule is not simply "pick the majority vote".

6. A binary symmetric channel receives as input a bit whose values $\{0, 1\}$ have probabilities $\{p, 1-p\}$, but in either case, a transmission error can occur with probability ϵ which flips the bit. The surface plot below describes the mutual information of this channel as a function $M(\epsilon, p)$ of these probabilities:



(a) At the point marked A, the error probability is $\epsilon = 0$. Why then is the channel mutual information minimal in this case: $M(\epsilon, p) = 0$?

(b) At the point marked B, an error always occurs ($\epsilon = 1$). Why then is the channel mutual information maximal in this case: $M(\epsilon, p) = 1$?

(c) At the point marked C, the input bit values are equiprobable ($p = 0.5$), so the symbol source has maximal entropy. Why then is the channel mutual information in this case $M(\epsilon, p) = 0$?