Aalto University School of Electrical Engineering
Department of Communications and Networking

EXAM
17.5.2011

S-72.2410 Information Theory

1. (6p.) Entropy. Consider the distribution

$X\backslash Y$	0	1	2
0	0	1/4	1/8
1	1/8	1/8	1/16
2	1/8	1/16	1/8

   (a) Find $H(X)$.
   (b) Find $H(Y)$.
   (c) Find $H(X|Y)$.
   (d) Find $H(Y|X)$.
   (e) Find $H(X,Y)$.
   (f) Find $I(X;Y)$.

2. (6p.) Terminology. Explain (in words, not mathematically)

   (a) broadcast channel,
   (b) channel encoder,
   (c) prefix code,
   (d) Markov chain,
   (e) universal code,
   (f) memoryless channel.

3. (6p.) Channels. Draw a channel (or argue why such a channel cannot exist)

   (a) (2p.) with two input values, two output values, and capacity 1 bit per transmission;
   (b) (2p.) with two input values, two output values, and capacity 0 bit per transmission; and
   (c) (2p.) with two input values, three output values, and capacity 1.2 bits per transmission.

   Hint: Recall that a channel has one node for each input value and one node for each output value. There are edges between the input side and the output side; these edges are labeled with probabilities. The probabilities related to the edges from a node on the input side must sum to 1.

4. (6p.) Source coding.

   (a) (1p.) Which codes give shorter expected length, Huffman codes or Shannon codes? (No motivation needed here.)
   (b) (3p.) Determine whether there exists a binary prefix code with codeword lengths $(l_1, l_2, l_3, l_4, l_5, l_6, l_7) = (2, 2, 3, 3, 3, 4, 4)$. If it exists, present the code; otherwise, give a nonexistence proof.
   (c) (2p.) It is a well-known fact that if $(l_1, l_2, \ldots, l_m)$ are optimal codeword lengths for a source distribution $(p_1, p_2, \ldots, p_m)$ and a q-ary alphabet, then the associated expected length L of an optimal code fulfills

$$H_q(X) \leq L < H_q(X) + 1.$$

   What trick can be used to reduce the overhead (+1) in the upper bound of the expected codeword length per input symbol when sending a sequence of many symbols?