Math 5251 Error-correcting codes and finite fields
Spring 2022, Vic Reiner
Midterm exam 1
Due Wednesday Feb. 23 by 11:59pm, via Canvas

Instructions: This is an open book, open library, open notes, open web, take-home exam, but you are *not* allowed to collaborate. The instructor is the only human source you are allowed to consult.

1. (25 points total; 5 points each) **True or False**. Your answers must be justified either by counterexamples or proofs to receive full credit.

(a) There exists a source W with $\#W = 5$ and some choice of word probabilities having a binary Huffman code with codewords of lengths $(1, 3, 3, 3, 3)$.

(b) There exists a prefix ternary encoding $f : W \to \{0, 1, 2\}^*$ for some eight word source W whose codewords have lengths

$$(\ell_1, \ell_2, \ell_3, \ell_4, \ell_5, \ell_6, \ell_7, \ell_8) = (1, 1, 2, 2, 3, 3, 3, 4).$$

(c) If one has a uniquely decipherable n-ary encoding $f : W \to \Sigma^*$ of a source W, so $\#\Sigma = n$, and every codeword $f(w_i)$ has length at most ℓ, then $\#W \leq n^\ell$.

(d) Assume that source $W = \{w_1, \ldots, w_m\}$ with strictly positive word probabilities $\{p_1, \ldots, p_m\}$ has some $p_{i_0} \geq \frac{1}{2}$. Then in any binary Huffman encoding of W, the length of the word encoding w_{i_0} will be 1.

(e) If source $W = \{w_1, \ldots, w_m\}$ with word probabilities $(p_1, \ldots, p_m)$ has a word $h(w_{i_0})$ of length 1 in one of its binary Huffman encodings $h : W \to \{0, 1\}^*$, then $p_{i_0} \geq \frac{1}{2}$.

2. Let the source $W = \{w_1, w_2, \dots, w_7\}$ have word probabiities

$$\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{16}, \frac{1}{16}\right)$$

(a) (10 points) Compute the (binary) entropy $H(W)$. An approximate decimal final answer is fine, but must be explained.

(b) (10 points) Compute the minimum among all uniquely decipherable binary encodings $f : W \to \{0,1\}^*$ of the average length of the code words $f(w_i)$.

3. (15 points) Let $W = \{w_1, \dots, w_m\}$ and $W' = \{w'_1, \dots, w'_{m'}\}$ be two memoryless sources, with word probabilities $(p_1, \dots, p_m)$ for W, and $(p'_1, \dots, p'_{m'})$ for W'. Define a new source $W \times W'$ whose words are ordered pairs (w_i, w'_j) with $w_i \in W$ and $w'_j \in W'$, and probabilities $P((w_i, w'_j)) = p_i \cdot p'_j$. Prove that

$$H(W \times W') = H(W) + H(W').$$

4. Suppose that we are sending length 6 binary words $w = b_1b_2b_3b_4b_5b_6$ with $b_i \in \{0,1\} = \mathbb{F}_2$ through a noisy binary symmetric channel (BSC) having error probability p for each bit sent.

(a) (5 points) Compute the probability of at least one error during transmission, as a function of p.

Now we choose to send w with two extra parity check bits as follows:

$$f(w) = b_1b_2b_3b_4b_5b_6b_7b_8$$

where

$$b_7 := b_1 + b_2 + b_3,$$
$$b_8 := b_4 + b_5 + b_6.$$

(b) (10 points) Compute the probability of at least one undetected error when w is sent as $f(w)$, again as a function of p.

(c) (5 points) Considering the image of f as a set of codewords $\mathcal{C}$ inside $\{0,1\}^*$ of length 8, what is the (binary) rate of the code $\mathcal{C}$?

5. (20 points) Prove by induction on m that, for any binary Huffman encoding of a source W of size m , the word lengths $(\ell_1, \dots, \ell_m)$ achieve equality in the Kraft-McMillan inequality, that is, $\sum_{i=1}^{m} \frac{1}{2^{\ell_i}} = 1$.