Math 5251 Error-correcting codes and finite fields
Fall 2021, Vic Reiner
Midterm exam 1
Due Wednesday Oct. 13 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. (30 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 word probabilities having a binary Huffman code with codewords of lengths $(2, 2, 2, 3, 4)$.

(b) There exists a prefix binary encoding $f : W \to \{0, 1\}^*$ for some source $W = \{w_1, w_2, w_3, w_4, w_5\}$ with codewords of length $(2, 2, 2, 3, 4)$.

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

(d) Any source W of cardinality $\#W = n^\ell$ with $n \geq 2$ and $\ell \geq 1$ has at least one uniquely decipherable encoding $f : W \to \Sigma^*$ in which all codewords $f(w)$ have length at most ℓ.

(e) Assume that source $W = \{w_1, \ldots, w_m\}$ with 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.

(f) A source $W = \{w_1, \ldots, w_m\}$ with word probabilities $\{p_1, \ldots, p_m\}$ that has a word of length 1 in one of its binary Huffman encodings must have some $p_{i_0} \geq \frac{1}{2}$.

2. (15 points) Assume one is encoding long binary words using a CRC with polynomial $g(x) = x^4 + x^3 + 1$ in $\mathbb{F}_2[x]$, that is, tacking on four 4 extra bits that represent the remainder upon division by $g(x)$. Prove that 2-bit errors that occur exactly 15 positions apart will go undetected.

3. Let source $W = \{w_1, w_2, \ldots, w_8\}$ have word probabiities

$$\left\{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \frac{1}{64}, \frac{1}{128}, \frac{1}{128}\right\}$$

(a) (5 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 among all uniquely decipherable binary encodings $f : W \to \{0,1\}^*$ of the average length of the code words $f(w_i)$.

4. Suppose that we are sending length 4 binary words $w = b_1b_2b_3b_4$ 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:

$$f(w) = b_1b_2b_3b_4b_5b_6$$

where

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

(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 6, 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 $(\ell_1, \ldots, \ell_m)$ achieve equality in the Kraft-McMillan inequality, that is, $\sum_{i=1}^{m} \frac{1}{2^{\ell_i}} = 1$.