

2nd homework assignment - Coding and compression

turn in by May 3rd, 2021.

Problem 1. For each of the following codes decide whether it is uniquely decodable. For each uniquely decodable code find an infinite string which can be decoded in two different ways if there is such a string. Show that prefix-free codes do not have such a sequence.

- a) $C_1 = \{110, 11, 10\}$, $C_2 = \{110, 11, 100, 00, 10\}$
- b) $C_3 = \{0, 10, 11\}$, $C_4 = \{0, 01, 11\}$
- c) $C_5 = \{0, 01, 10\}$, $C_6 = \{0, 01\}$

Problem 2. Consider prefix-free code over binary alphabet with codeword lengths $\ell_1, \ell_2, \dots, \ell_k$ such that

$$\sum_{i=1}^k 2^{-\ell_i} < 1.$$

Show that there are arbitrary long strings in $\{0, 1\}^*$ that cannot be split into a sequence of codewords.

Problem 3. Consider a set of integers $x_1 < \dots < x_n$ with associated probabilities p_1, p_2, \dots, p_n . Consider a static (fixed) binary search tree for the set of these items. Show a lower bound in the form of entropy on the expected length of a search in the tree, where item x_i is searched for with probability p_i . How does the answer change if we allow a dynamic tree that can be reorganized after each search.

Problem 4. Give an example of a random variable X and an arbitrary code $C : X \rightarrow \{0, 1\}^*$ such that $E[|C(X)|] < H(X)$. For any random variable X and a code C , show a lower bound on $E[|C(X)|]$ in terms of $H(X)$ which is as good as you can come-up with.