Information Theory
Problem Set 04 - Symbol Codes

Luís Felipe Ramos Ferreira

lframos_ferreira@outlook.com

1. (a) A (binary) symbol code for an ensemble, denoted by C, is a function that maps the outcomes of the ensemble to a set os binary strings. In particular, this set of strings is a subset of $\{0, 1\}^+$, which denotes the set of all binary strings of non zero length. The extended code for the ensemble, denoted by C^+, is a function from $\mathcal{A}_X^+$ to $\{0, 1\}^+$. More precisely, it represents the concatenation of the codewords of a ordered set of outcomes from the ensemble.

   (b) A symbol code is uniquely decodeable when no element is mapped to the same codeword. It is easy to see that is true based on the pidgeon-hole principle. More formally, a code $C(x)$ is uniquely decodeable if, under the extended code C^+, we have:

$$\forall x, y \in \mathcal{A}_X^+, x \neq y \Rightarrow c^+(x) \neq c^+(y)$$

   A symbol code is prefix-free if no codeword is a prefix of any other codeword, as stated by McKay [1].

   (c) c

   (d) d

2. 2

3. 3

   4.

5. 5

6. 6

References

[1] David J. C. MacKay. *Information Theory, Inference and Learning Algorithms.* 7th edition, 2005.