

The University of Oxford

MSc (Mathematics and Foundations of Computer Science)

## Information Theory

Michaelmas Term 2023

*Students should complete ONE of the two possible projects below.*

### Project 1

Consider the Reed–Solomon code, which extends the Hamming error correcting code considered in the course. Write a report which details:

- the construction of the code, from the perspective of polynomials in finite fields,
- the process of encoding and decoding using the code,
- gives examples (based on your own implementation of the code) of encoding and decoding, and
- demonstrates the error correction capabilities of the approach.

Your report should include sufficient detail for (at least one interpretation of the) algorithm to be reproduced in full. Code should be submitted in an appendix, and may make use of standard libraries for finite-field arithmetic (but clearly may not use a prior implementation of the Reed–Solomon code!). You will need to find and consult appropriate references, the wikipedia page on Reed–Solomon codes may be a good place to start, but should not be your only reference.

### Project 2

Consider the Lempel–Ziv–Welch data compression code, which is an alternative to arithmetic and Huffman coding. Give a summary and your own implementation of the classic fixed-width LZW code (code submitted in an appendix). Write a report which details:

- the construction of the code,
- the process of encoding and decoding using the code,

- gives examples (based on your own implementation of the code) of encoding and decoding, and
- compares the compression performance of LZW, Huffman and BAC algorithms (including the one-step Markov chain version of BAC, where probabilities can be estimated from the file, and stored as an initial header to the encoded data).

Your report should include sufficient detail for (at least one interpretation of the) algorithm to be reproduced in full. For testing, you should use a range of moderately sized uncompressed files (e.g. texts from project Gutenberg (stored as .txt), data files from ONS or similar (stored as .csv)). You will need to find and consult appropriate references, the wikipedia page on LZW codes may be a good place to start (as is Welch's 1984 paper)..