

Student: Ty Davis

Course: ECE 5420

Subject: Programming Assignment 2 - LZ Compression Algorithm

Date: September 21, 2025



1 The entropy of the source

There are 100 0's and 900 1's in the `source1.txt` file. We will use the equation for entropy listed below:

$$H(X) = -\sum_{i=1}^M p_i \log_2(p_i)$$

We find that

$$H(X) = -0.1 \log_2(0.1) - 0.9 \log_2(0.9) = 0.469$$

2 How many bits are used on average to represent each information bit after compression?

For the `source1.txt` file, the resulting information was 840 bits long, so the average compression rate is 0.84.

3 Is it possible to find a better compression algorithm? Why or why not?

Yes, the entropy represents the “theoretical lower bound of how many bits are required to encode a source symbol.” So there is, in theory, a lossless compression of this data that will take only ≈ 470 bits.

4 What's your W number?

W01367741