ENEE5304, INFORMATION AND CODING THEORY

Course Project on Source Coding

Due: July 1, 2023 (via ITC)

Course Project

The Written Report

- Double space, 12-point font.
- At least two recent references.
- Write the report in your own words. Do not just copy and paste. If you quote something, cite the reference
- Sections: Define the problem in the introduction, Method (or theoretical background), Results (or Simulations or implementation) and their analysis, the code (appendix), Conclusions, and References.

Presentation

- Students will be required to present their work in my office at designated dates, to be announced later.

Course Project on Lempel-Ziv Encoding of Random Symbols

1. Generate a random sequence of symbols a, b, c, d, such that P(a)=0.4, P(b)=03, P(c) = 0.2, P(d)=0.1
2. Calculate the source entropy H in bits.
3. First, let the size of the sequence be **N=20 symbols**.
4. Develop a program to parse the data and assign a number to each phrase. **Submit the result in your report**
5. Find the number of binary digits $\mathbf{N_B}$ needed to encode the sequence and the **number of bits per symbol** assuming that the tail of each codeword is in ASCII format (8 bits). **Submit the result in your report**
6. Repeat parts 4 and 5 five times and find the average value of ($\mathbf{N_B}$)
7. Find the compression ratio relative to the ASCII code (average value of $\mathbf{N_B}$ /N*8)

Course Project on Lempel-Ziv Encoding of Random Symbols

Repeat your calculations and fill in the following table

Sequence length N	Size of encoded sequence (N_B)	Compression ratio N_B /(8*N)	Number of bits per symbol (N_B /N)
20			
50			
100			
200			
400			
800			
1000			
2000			

Course Project 1 on Huffman Code

- Use Huffman code to find the codewords for the characters (a, b, c, d)
- Find the average number of bits/character for the code.
- If ASCII code is used, find the number of bits needed to encode 100 randomly generated symbols.
- Make the comparison between the Huffman and Limpel-Ziv codes for the case when N=100 as shown in the table below.

Sequence length N	Size of encoded sequence (N_B)	Compression ratio N_B /(8*N)	Number of bits per symbol (N_B /N)
100 (Limpel-Ziv)			
100 (Huffman)			