

# **ENEE5304, INFORMATION AND CODING THEORY**

## **Course Assignment on Source Coding**

**Due: June 15, 2024 (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 via zoom at designated dates, to be announced later.

## Course Project on Lempel-Ziv Encoding of Random Symbols

- Generate a random sequence of symbols a, b, c, d, such that  $P(a)=0.5$ ,  $P(b)=0.3$ ,  $P(c) = 0.1$ ,  $P(d)=0.1$
- Calculate the source entropy  $H$  in bits.

The following part is to be done in complete details

- First, let the size of the sequence be  **$N=30$  symbols**.
- Develop a program to parse the symbol sequence and assign a number to each phrase.

**Submit the result of the simulation in your report**

- Find the number of binary digits  $N_b$  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**

## Course Project on Lempel-Ziv Encoding of Random Symbols

- Find the compression ratio relative to the ASCII code ( $N_B / N * 8$ ).
- Summarize your results in the 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			

# Course Project on Lempel-Ziv Encoding of Random Symbols

Repeat parts 4 and 5 five times and find the average value of ( $N_B$ ), call it  $N_{av}$   
Repeat your calculations for various values on N and fill in the following table

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

## Comparison with 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 Lempel-Ziv codes for the case when  $N=100$  as shown in the table below.

Sequence length $N$	Size of encoded sequence ( $N_{av}$ )	Compression ratio $N_{av} / (8*N)$	Number of bits per symbol ( $N_{av} / N$ )
100 (Lempel-Ziv)			
100 (Huffman)			