

COMP9319 Exercises

Answers : To be released one week later.

Question 1

Given the text string below:

jejunojejunostomy

- What is its entropy?
- Draw a Huffman tree based on the letters and their corresponding distributions for the above text string (Do not need to draw trees for the intermediate steps).
- Provide the resulting Huffman code for each letter.
- What is the average number of bits needed for each letter, using your Huffman code? How does it compare to the entropy ? (i.e., equal/larger/small and why)

Question 2

- The length of a given string is 8, containing letters a, f, i, r with their probability ranges as below:

a [0.0, 0.125), f [0.125, 0.625), i [0.625, 0.75), r [0.75, 1.0)

Decode the arithmetic code 0.91805 to its corresponding string.

- Given the string:

jejunojeju

Derive an arithmetic code. (Your answer should be in decimal number with minimum precision).

Question 3

Consider the dictionary-based LZW compression algorithm. Suppose the alphabet is the set of ASCII characters, and the first 256 (i.e., <0> to <255>) table entries are initialized to these characters.

Show the dictionary (symbol sets plus associated codes) and output for LZW compression of the input string:

jejunojejuno