

Data Structures and Algorithms – (COMP SCI 2C03)  
Winter 2021  
Tutorial - 9

April 5, 2021

1. How many letter comparisons would KMP algorithm perform on the text  $a^n$  and pattern  $a^{m-1}b$ .
2. Compute the border array of the string `abaababaabaab`
3. Compute the border array of the string  $w = a^n$ .
4. Consider the four variable-length codes shown in Figure 1. Which of the codes are prefix-free? Which of the codes are uniquely decodable? For those that are uniquely decodable, give the encoding of 10000000000000.

symbol	code 1	code 2	code 3	code 4
A	0	0	1	1
B	100	1	01	01
C	10	00	001	001
D	11	11	0001	000

Figure 1: Table for Question 4

5. How many bits are needed to encode  $N$  copies of the symbol `a`, as a function of  $N$ ) using run-length encoding? How many bits are needed to

encode  $N$  copies of `abc` (as a function of  $N$ ) using run-length encoding? (you may consider ASCII encoding)

6. Give the result of encoding the strings `a`, `aa`, `aaa`, `aaaa`, ... (strings consisting of  $N$  a's) with run-length and Huffman encoding.
7. Compute the bitstream encoding of the binary trie given in Figure 2 representing prefix-free codes for letters A, B, C, D, E.

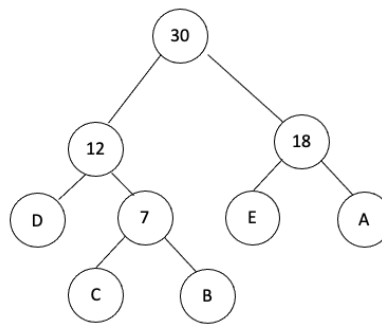


Figure 2: Prefix-free code binary trie

8. Consider the input string  $w = abababacaaaadaeaceeaabbb$ . Encoding  $w$  in 8-bit ASCII requires how many bits? How much savings in terms of bits is achieved if  $w$  is encoded using Huffman encoding (remember to include the number of bits required to encode the binary trie)?