UNIVERSITY OF LIVERPOOL

Department of Computer Science
Sebastian Wild

Date: 2020-04-22
Version: 2020-05-10 22:34

Tutorial 8 for COMP 526 – Applied Algorithmics, Winter 2020

—including solutions—

It is highly recommended that you first try to solve the problems on your own before consulting the sample solutions provided below.

Problem 1 (Huffman code)

Compress the text $T =$ `HANNAHBANSBANANASMAN` using a Huffman code; give

1. the character frequencies,
2. a step-by-step construction of the Huffman tree,
3. the Huffman code, and
4. the encoded text.
5. Finally, compute the compression ratio of the result.

Solutions for Problem 1 (Huffman code)

1. Frequencies:

char	A	B	H	M	N	S
weight	7	2	2	1	6	2

2. We proceed step by step; remember the tie-breaking rules!

   (i) mins: `M` and `B`, new symbol `MB` with weight 3.

char	A	MB	H	N	S
weight	7	3	2	6	2

   (ii) mins: `H` and `S`, new symbol `HS` with weight 4.

char	A	MB	HS	N
weight	7	3	4	6

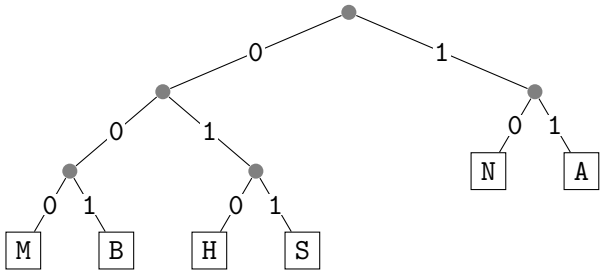
(iii) mins: MB and HS; new symbol with weight 7

char	A	MB HS	N
weight	7	7	6

(iv) min: N and A, new symbol NA with weight 13

(v) only two left, new symbol MB HS NA

Note that the nested boxes encode the trie uniquely:



3. The actual Huffman code is

char	A	B	H	M	N	S
code	11	001	010	000	10	011

4. $C =$ 010 11 10 10 11 010 001 11 10 011 001 11 10 11 10 11 011 000 11 10

5. Compression ratio is

$$\frac{46}{20 \cdot \lg(6)} \approx 0.890$$

Problem 2 (Hamming code)

We consider the $4 + 3$ Hamming code from class.

1. Given the message 0101, determine the parity bits and the final transmitted block.
2. Is 1111111 a valid block, i.e., have (detectable) errors occurred?

Solutions for Problem 2 (Hamming code)

1. parity bits 101, block 0101101.
2. Yes, it is a valid block: the checking bits are 0.

This implies that no 1-bit error has occurred; we cannot know whether 3 or more bit have flipped simultaneously, though.