

01.1 Explain what the term data compression means.

[2 marks]

Amount of computer memory [1] required to store the data [1] is reduced [1]

Vague answers such as "makes it smaller" or "makes files smaller" – 1 max

01.2 Explain why it might be desirable to use data compression.

[4 marks]

Explain, so look for developed points

Amount of memory needed to store files is reduced [1] so more files can be stored [1] less memory is needed [1] which is cheaper [1]

Time taken to transfer files [1] is reduced because the files are smaller in Bytes [1]

Time taken to render images [1] or download media files [1] from websites is reduced [1]

02 Run length encoding (RLE) is a method of data compression which uses frequency/data pairs.

02.1 State a run length encoding of the series of characters ttjjeeess

[1 mark]

2t2j3e2s

02.2 Explain how run length encoding (RLE) uses frequency/data pairs to compress data

[2 marks]

The number of each data item is recorded rather than recording all the data items [1] so repeated data can be stored using less memory [1]

02.2 Explain why run length encoding (RLE) does not always result in effective data compression.

[2 marks]

Single frequency data items [1] take more data to encode [1] – credit examples (e.g. trgt would be encoded as 1t1r1g1t which is longer)

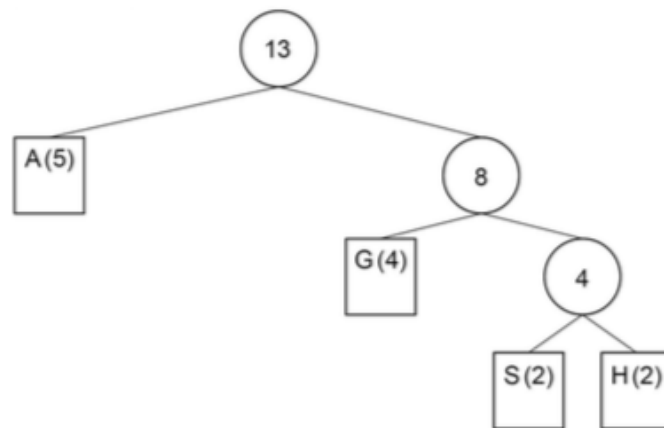
If there isn't lots of repeated data [1] it will take more data to encode [1]

03 Huffman coding is another method of data compression used to compress text. The Huffman tree in **Figure 2** was created to encode the string shown in **Figure 1**.

Figure 1

AAGHHGGSAAASG

Figure 2



03.1 Complete the code table below for characters G, S and H for the Huffman tree shown in Figure 6. The code for character A has already been completed.

[3 marks]

Character	Binary code
A	0
G	10
S	110
H	111

03.2 Explain how data is compressed using Huffman Coding. You do not need to explain how a Huffman Tree is created.

[3 marks]

More frequently occurring letters [1] are put at the top of the tree [1]. This means that they require less data to encode [1]. Because these occur more often the amount of data needed will be reduced [1]

Credit comparison with using ASCII code (7 bits) for 2 max unless idea of most frequent at top so require less data included.