3.3.8 Data compression (Huffman coding)

Teacher notes

Key information

It is reasonable to say that the Huffman method of encoding takes a little more effort to carry out than the Run Length Encoding (RLE) method discussed previously.

Reasons for this are:

- each encoding generates its own unique tree
- drawing the diagrams can be time-consuming, regardless of the media used, and it is easy to make a mistake in the process'
- the greater the variety of characters, the greater the complexity of the tree generated.

In general, most examples available elsewhere are extremely simplistic, eg they use only a handful of the different ASCII characters available and ignore the character case (which is unrealistic).

It is probably worth pointing out to students that the example provided in the PowerPoint is of a middle level.

Again, it may well be worth providing the students with the sheets containing the development of the Huffman tree, as it is likely to be challenging for students who need additional support.

Useful references

The following website gives a good practical explanation of the Huffman coding method without being over-complicated. It would provide useful background reading for a teacher – the worked examples are very instructive.

algoviz.org/OpenDSA/Books/Everything/html/Huffman.html#