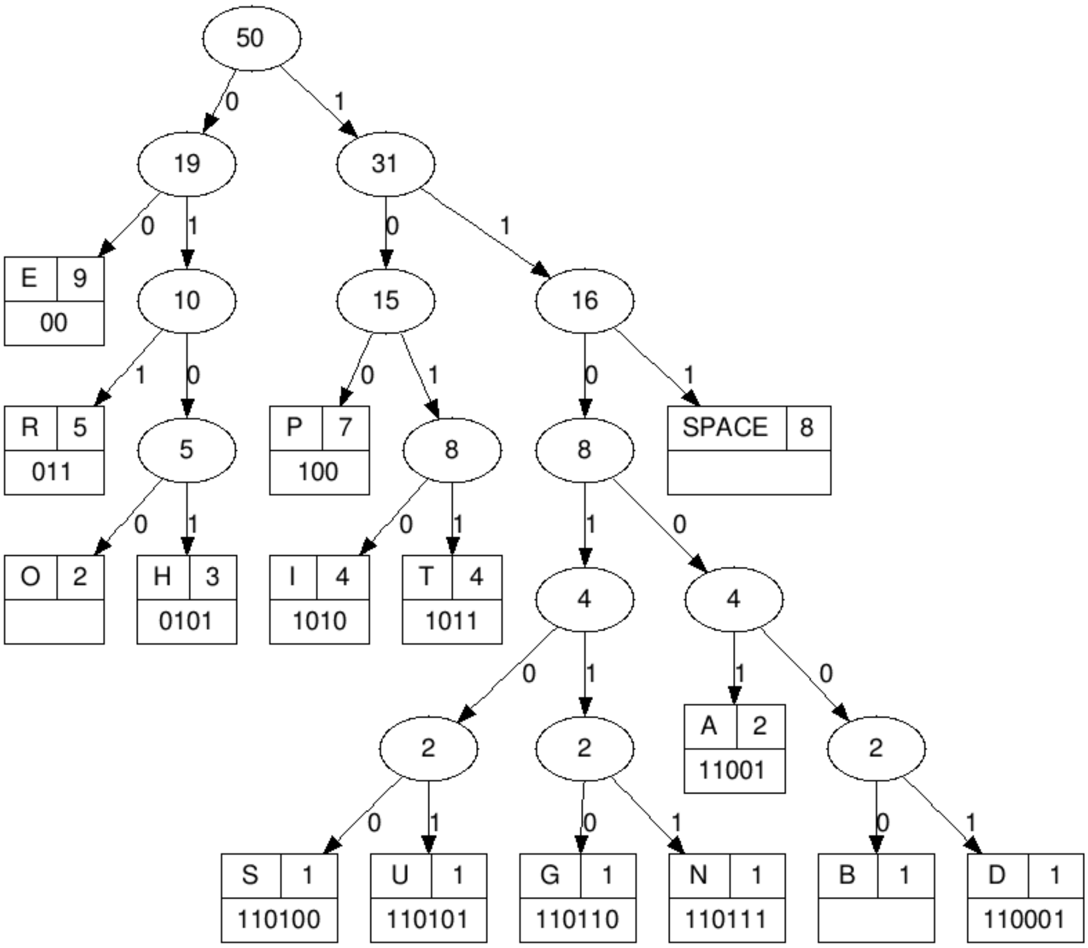
Representations of Data

1. This question is about file sizes.
    a. Calculate the file size in *bits* for a five-minute sound recording that has a sample rate of 96 kilohertz (kHz) and a sample resolution of 5 bits. You should show your calculations.
    b. Another sound file has a size of 48,000 bits. What is 48,000 bits in kilobytes? You should show your calculations.
    c. Sound files are stored as bit patterns. Bit patterns are often compressed. Compress the following bit pattern using run length encoding: 1111 1100 1111 1000 0000 1111 1111 0001
    d. Which of these statements is always true about Run Length Encoding?
        i. It will always make a file smaller.
        ii. It is most effective on data that appears random.
        iii. It will not lose any of the original data

2. This question is about number bases. A certain bit pattern is as follows: 1010 0111
    a. Convert the bit pattern into decimal.
    b. Convert the bit pattern into hexadecimal
    c. A student's answer to the question "Why is hexadecimal often used instead of binary?" is "*Because it uses fewer digits it will take up less space in a computer's memory.*" Explain why the student's answer is incorrect.
    d. Why do we use hexadecimal? Give one correct reason, with an explanation.
    e. Explain how a binary number can be multiplied only by shifting bits.
    f. What shift is needed to multiply a binary number by 16?
    g. What does a single right-shift do to a decimal number?

3. This question is about character encoding. ASCII (American Standard Code for Information Interchange) is a coding system that can be used to represent characters using 7 bits.
    a. In ASCII the character A (*uppercase-A)* is represented by the numeric code 65. What character is represented by the numeric code 75?
    b. Unicode is an alternative to the ASCII coding system. Describe one advantage and one disadvantage of using Unicode to represent characters instead of using ASCII.
    c. Charlie is using his computer to program an LCD display embedded into a piggy bank he has designed, which counts the coins inserted. He wants it to show the balance when a button is pressed, displaying the message: "You have currently stored £X.XX". He notices when testing that the output is garbled: "You have currently stored ??X.XX". The embedded computer that drives the display is very basic, and only supports ASCII. Can you explain what might have happened?

4. This question is about data compression. When data is stored in a computer it is often compressed. One method that can be used to compress text data is Huffman coding. To produce a Huffman code each character in a piece of text is placed in a tree, with its position in the tree determined by how often the character was used in the piece of text. A Huffman tree for the phrase `PETER IS THE PIPER IN THE PETERBOROUGH PIPE PARADE` is shown below, with some of the codes missing:



Using this tree, the letter G would be represented as 110110, and the letter R as 011.

a. Using the Huffman tree in Figure 3, determine the Huffman coding for the characters O, SPACE and B, all of which are missing.
b. Using Huffman coding the phrase above, including spaces, can be stored in 176 bits. Calculate how many additional bits are needed to store the same piece of text using standard ASCII. Approximately what percentage of the uncompressed text is the Huffman coded version? Show your working.

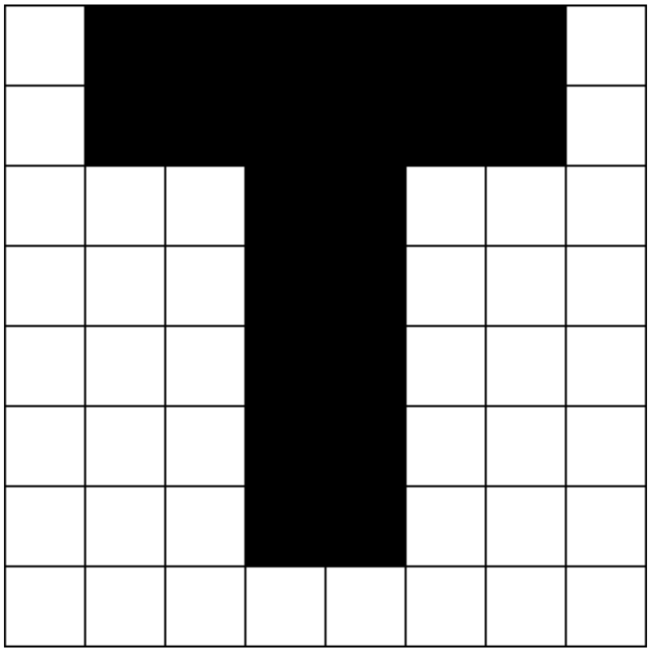


Figure 1

5. This question is about images. Figure 1 is an image depicting the letter “T”, with 64 pixels.
    a. What is a pixel?
    b. Explain why this image can be stored using 64 bits.
    c. How many bits would be needed if the image used eight colours?
    d. Express the image using Run Length Encoding. Use W for white, and B for black. Start each row on a new line.
    e. Explain the term *image resolution*, and explain how changing it affects the file size.

6. This question is about binary arithmetic.

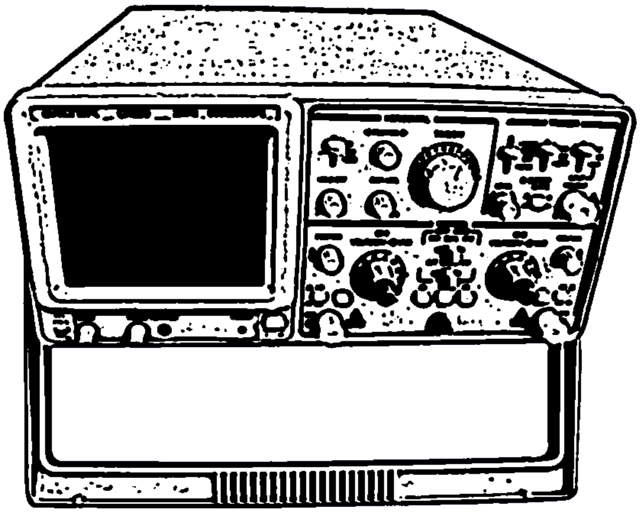
    a. What is 0011 1001 + 0101 1010?

    b. What is 0110 1010 + 0111 1011?

    c. What is 0111 1111 + 1110 0111?

    d. If the addition in part c. was performed on an 8-bit CPU, what would happen? What do we call this?

7. This question is about sound.



The device shown above is known as an *oscilloscope*, and the screen on the left shows a visual image of a signal or sound wave sent into its input. Eshaal is using it to view the sound waves generated by a microphone placed underneath a tank of fish.

Eshaal is running an audio recording program on her computer, which will record sound detected by the microphone. She plans to record the sound of her fish tank over a one hour period. She sets the program to record at a *Sample Resolution* of 32 bits, and a *Sample Rate* of 1MHz.

a. Describe the features of the sound waves that would be seen on the screen. Include a diagram if it makes your explanation clearer.

b. Computers cannot process sound waves directly, because they are analogue. Describe the process of converting analogue sound to digital data.

c. Are Eshaal's settings for her recording sensible? Why? Why not?

d. Eshaal wants to be able to send a five-minute section of the recording via email, with a maximum file size of 25 Megabytes. The program only supports a *sampling resolution* of 32 bits. To the nearest 100Hz, what is the highest sampling rate she can select? Show your calculations.