Data Representation

Keyword	Definition
Bit	A single 0 or 1
Byte	8 bits
Decimal	Base 10 number system
Binary	Base 2 number system
Hexadecimal	Base 16 number system (Hex)
ASCII & Unicode	Character sets that can be represented by a computer
Metadata	Data about data
Lossy compression	Permanently loses some data when compressed
Lossless compression	Will not lose any of the original data when compressed

A sense of scale

File	Size
One character of text	1 byte
A full page of text	30 kB
One small digital colour photograph	3 MB
Music CD capacity	650 MB
DVD capacity	4.5 GB
Hard disk capacity	1 TB

Bits and Bytes

- 1 bit = a single 0 or 1 (**b** represents bit)
- 1 byte = 8 bits = 1 character of text (**B** represents byte)
- (A nibble = 4 bits or half a byte)
- 1 kB (kilobyte) = 1000 bytes
- 1 MB (megabyte) = 1000 kB or 1000x1000 bytes
- 1 GB (gigabyte) = 1000 MB
- 1 TB (terabyte) = 1000 GB

Most and least significant bit

- The most significant bit (MSB) is the bit with the largest value
  - This is the bit that is furthest to the left
- The least significant bit (LSB) is the bit with the smallest value
  - This is the bit that is furthest to the right

128	64	32	16	8	4	2	1
1	0	0	1	1	1	0	1

Hex to Decimal conversion

16s		Units	
2		A	
2x16	+	10	= 42 in decimal

Multiply the left-hand digit by 16, then add the units

Hexadecimal

Decimal	Hex	Decimal	Hex
1	1	9	9
2	2	10	A
3	3	11	B
4	4	12	C
5	5	13	D
6	6	14	E
7	7	15	F
8	8	16	10

Binary representation

Number of Bits	Possible states
1	2
2	4
3	8
4	16
5	32
6	64
7	128
8	256

Decimal to Binary

- How do you convert 28 to binary?
- Method
  - Working right to left, write out the numbers 1, 2, 4, 8 and so on, doubling each time to 128

128	64	32	16	8	4	2	1
0	0	0	1	1	1	0	0

- 128, 64 and 32 are all greater than 28, so put a zero for these
- Put a 1 in the 16 column, 28-16=12
- Put a 1 in the 8 column, 12-8 = 4
- Put a 1 in the 4 column, 4-4=0 so put zero in other columns

Data Representation

Decimal to Hex conversion

1. Divide the denary number by 16 and record the remainder. — 164 ÷ 16 = 10 remainder 4
2. The remainder, 4, is the number of 1s, and the answer, 10, is the number of 16s. — 10 in hex is A

16	1
A	4

164 is equivalent to A4.

You can also convert denary to hex by converting denary to binary digits and then converting binary digits to hex.

Binary to Hex conversion

1. Split 01001100 into nibbles, and place each nibble into a base 2 place value table.

0100				1100			
8	4	2	1	8	4	2	1
0	1	0	0	1	1	0	0

2. For each nibble, add the column headings for the columns that contain a 1.

4   8 + 4 = 12

3. Convert the denary values to hex.

4 = 4   12 = C

01001100 is equivalent to 4C.

Hex to Binary conversion

1. Convert each hex digit to binary.

A				6			
8	4	2	1	8	4	2	1
1	0	1	0	0	1	1	0

A = 10 = 8 + 2   6 = 4 + 2

2. Combine the nibbles to make the binary number.

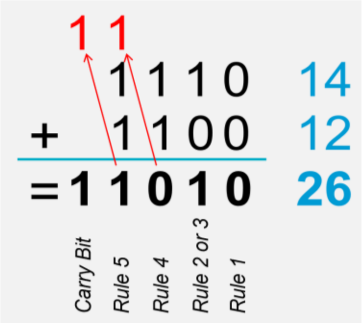
10100110

A6 is equivalent to 10100110.

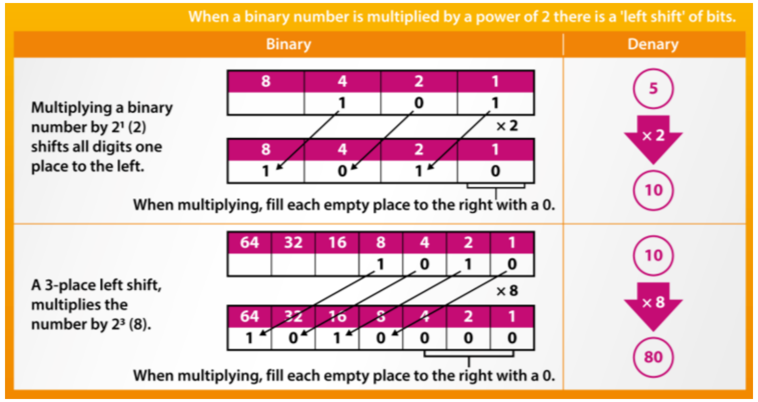
Binary Addition

Work right to left and apply these simple rules:

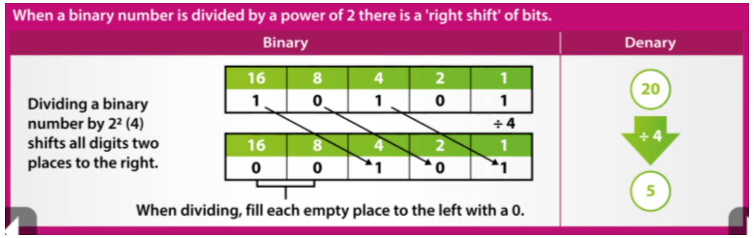
1. 0 + 0 = 0
2. 0 + 1 = 1
3. 1 + 0 = 1
4. 1 + 1 = 0 *Carry 1*
5. 1 + 1 + 1 = 1 *Carry 1*



Binary Shifts - Multiplication



Binary Shifts - Division



Calculating image size

Size in bits = Width x Height x Colour depth

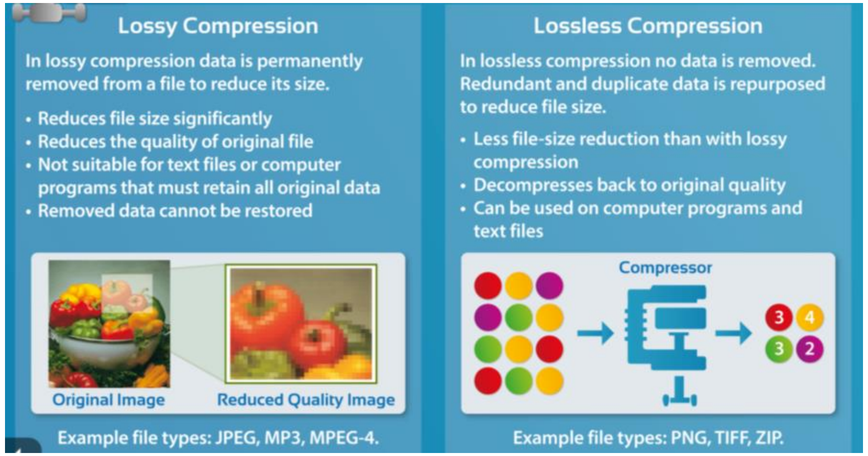
Size in bytes = Width x Height x Colour depth / 8

Calculating sound file size

Size in bits = Sampling Rate x Resolution x Duration

Size in bytes = Sampling Rate x Resolution x Duration / 8

Data Compression



RLE – Run Length Encoding

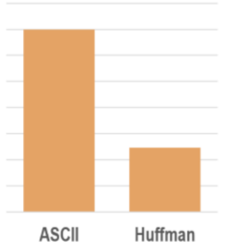
Assuming 1 = white and black = 0, this run could be encoded as:

1111000000000000

Using RLE, the row could be represented as 4 '1's and 12 '0's or 4 1 12 0

Huffman compression

- Huffman coding of "*she sells sea shells*" = 49 bits
- 7-bit ASCII coding = 20 Characters or 140 bits
- This is a lossless reduction in the file size of 65%



Huffman Trees

Consider the frequency of each character in the sentence: "SHE SELLS SEA SHELLS"

Character	Frequency
S	6
E	4
L	4
Space	3
H	2
A	1