Games of Computer Images and Data Encoding

Key Stage: 3

Strand: Number and Algebra Strand and Data Handling Strand

Learning Unit: Laws of integral indices
Measures of central tendency

Objective:
(i) To apply knowledge of binary numbers to interconvert between text and image and digital data
(ii) To select the appropriate text coding method
(iii) To apply mathematics in tackling daily-life problems

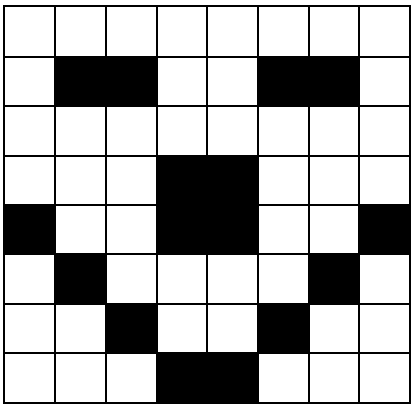
Pre-requisite knowledge:
(i) Interconversion between binary numbers and denary numbers
(ii) Finding mean, median and mode of ungrouped data
(iii) Recognising the concept of weighted mean

Relationship with other KLA(s) in STEM Education:
Topic C "Data Representation" in The Compulsory Part of Technology Education Key Learning Area Information and Communication Technology Curriculum and Assessment Guide (Secondary 4 - 6)

Description of the Activities:

Warm-up activity 1

The teacher displays the image with a human face (as shown in the right figure) to students. There are 8 rows of squares and each row consists of 8 squares.



Questions for discussion:

1. If we represent a white square by a digit 0, and a black square by a digit 1, how can we represent the image at the right side by binary numbers?

Notes for Teachers:

(i) Counting from the left upper corner and going downwards row by row, we have 8 binary numbers each of 8 digits: 00000000, 01100110, 00000000, 00011000, 10011001, 01000010, 00100100 and 00011000

(ii) Teachers may assign students to convert the binary numbers to denary numbers for each row, take the second row as an example:

$$01100110 = 2 \times 1 + 2^2 \times 1 + 2^5 \times 1 + 2^6 \times 1$$

$$= 2 + 4 + 32 + 64$$

$$= 102$$

(iii) After several calculations, students could use a list of eight denary numbers to represent the human face image.

Activity 1

1. Divide students into groups of two to four students. Each group will have two sheets of 8 × 8 squares graph paper (1 Graph Paper I & 1 Graph Paper II)
2. After discussion, each group will design and draw an image on Graph Paper I secretly.
3. Students use the method used in the warm-up activity to transform the image into a list of binary numbers.
4. Each group then converts the eight binary numbers into denary numbers.
5. The denary numbers are then written on the blanks of Graph Paper II.
6. Each group hands out their Graph Paper II with denary numbers.
7. Teachers collect all the Graph Paper II and distribute them to different groups. Each group needs to rebuild the image according to the given denary numbers.
8. The first group who can rebuild the correct image will get a prize.

Questions for discussion:

1. If one group does not follow the teacher's instruction strictly and represents a white square by a digit 1 and a black square by a digit 0, do you have any method to redraw its image quickly?
2. If one group forgets to write one of the eight denary numbers (the row number is known), at most how many times does another group of students need to try to get the correct image by the trial-and-error method?

Notes for Teachers:

1. Transfer the denary numbers to binary numbers, change the digits from 1 to 0 and 0 to 1 and then redraw the figure. Students may also try to subtract the denary numbers from 255 (i.e. 11111111 in binary) and then convert the result to binary numbers for drawing the figure.

2. The number is 256 as from 00000000 to 11111111 (i.e. 0 to 255), there are 256 different binary numbers.

Warm-up activity 2

We know that computers process text and images in a binary system. The length of a string of binary code representing text or images affects the cost of data storage and transmission. To reduce costs, we should shorten the number of bits required to send the same information. We should use the least-symbol encoding method. This activity is to compare different methods of encoding and to get the better choice.

1. The teacher intends to transfer the message "KEN DELETES THE FILE" to his friend by a computer. For the 10 different letters and the space in the message, if the computer uses 4 bits to encode the message, one of the method is as follows:

Alphabet	K	E	N	D	L	T
Code	0000	0001	0010	0011	0100	0101

Alphabet	S	H	F	I	Space
Code	0110	0111	1000	1001	1010

The teacher may point out that using this encoding method, out of the 16 combinations for 4 bits of code, we only use 11 of them to represent the message. Students are asked to suggest more efficient coding methods.

2. The teacher shows another encoding method using 3 bits originally, and using 4 bits for the rest:

Alphabet	K	E	N	D	L	T
Code	000	001	010	011	100	101

Alphabet	S	H	F	I	Space
Code	110	111	0000	0001	0010

Students are required to calculate the average length of the code for each alphabet.

Notes for Teachers:

(i) If we encode the message as in (2) above, the average length of the code for each alphabet

$$= \frac{3+3+3+3+3+3+3+3+4+4+4}{11}$$

$$= \frac{8(3)+3(4)}{11}$$

$$= 3.27$$

(ii) Encoding method in (2) has a disadvantage. According to this encoding method, take 0000001101 as an example, the computer may recognise the message as KIT, or it may recognise the message as FET. It is impractical to have two meanings for the same code.

Activity 2

1. Teacher introduces the following two encoding methods to students:

   Method 1:

Alphabet	D	E	L	T	H	F
Code	000	1100	010	011	100	101

Alphabet	I	Space
Code	001	1101

   Method 2:

Alphabet	D	E	L	T	H	F
Code	11110	0	110	1110	111110	1111110

Alphabet	I	Space
Code	11111110	10

2. Students form groups of two to four. According to the following conditions, students need to judge which encoding method is better:

   (i) Consider the average length of the code for each alphabet

   (ii) When they are used for encoding the message "DELETE THE FILE"

3. One student representative of each group explains their choice.

Questions for discussion:

1. The frequency for the occurrence of alphabets is not even in normal state. Do you know which alphabet appears the most? Which alphabet appears the least?
2. Based on the frequency of the occurrence of different alphabets, can you design a better encoding method? Try the method out and make explanation.

Notes for teachers:

(i) Unlike the coding method in the Warm-up Activity 2, messages encoded by Method 1 and Method 2 have only one way of interpretation.

(ii) Suggested solutions for question 2 (i) and (ii) in Activity 2:

Question 2(i):

By Method 1, the average length of the code for each alphabet

$$= \frac{3+4+3+3+3+3+3+4}{8}$$

$$= \frac{6(3)+2(4)}{8}$$

$$= 3.25$$

By Method 2, the average length of the code for each alphabet

$$= \frac{5+1+3+4+6+7+8+2}{8}$$

$$= \frac{36}{8}$$

$$= 4.5$$

The length of a string of codes affects the cost of storing and transmitting data. Method 1 is better if the average length of the code for each alphabet is considered.

Question 2(ii):

The total number of bits needed for encoding the message by Method 1 is 52, the average length of the code for each alphabet

$$= \frac{3+4+3+4+3+4+4+3+3+4+4+3+3+3+4}{15}$$

$$= \frac{8(3)+7(4)}{15}$$

$$= 3.47$$

The total number of bits needed for encoding the message by Method 2 is 49, the average length of the code for each alphabet

$$= \frac{5+1+3+1+4+1+2+4+6+1+2+7+8+3+1}{15}$$

$$= \frac{49}{15}$$

$$= 3.27$$

For this specific message, Method 2 is better as alphabet *E* occurs many times in the message. We need 4 bits to encode the alphabet *E* by Method 1, but we only need 1 bit to encode it by Method 2. Hence, it saves many storage spaces.

(iii) Answers for "Questions for discussion" of Activity 2

1. Refer to https://www.math.cornell.edu/~mec/2003-2004/cryptography/subs/frequencies.html, *E* appears most frequently, which is 12.02% based on a sample of 40,000 words and *Z* appears least frequently, which is just 0.07%.

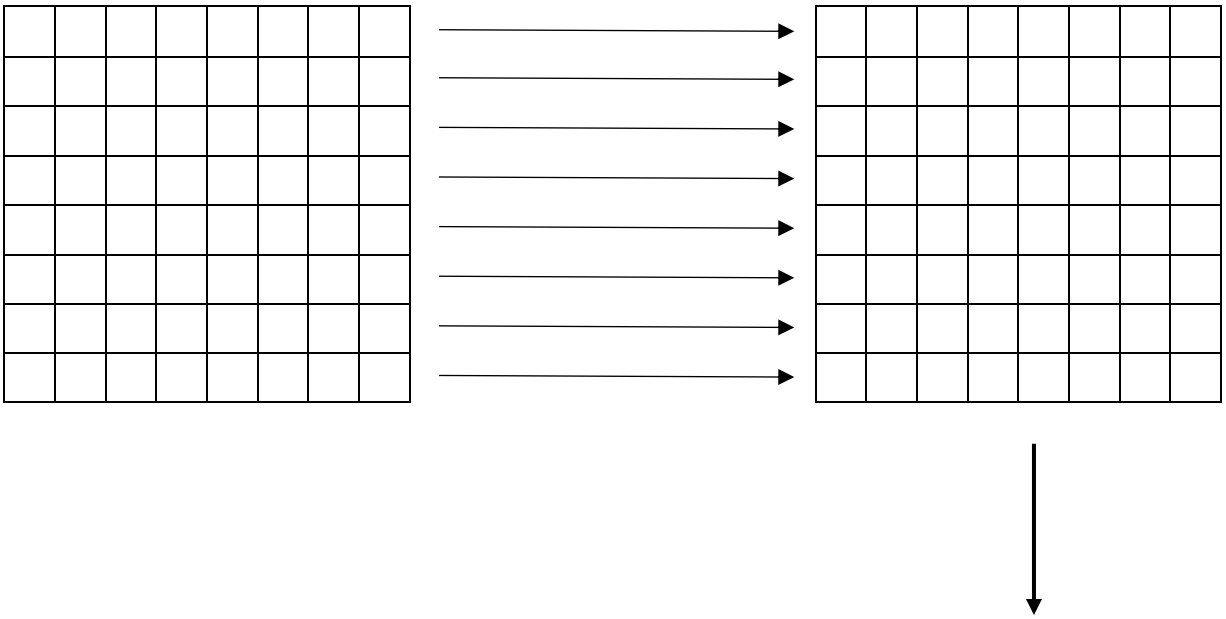
2. We use small number of bits to encode the alphabets *E*, *T*, *A*, *O*, *I* and *N*, which have higher frequency of appearance. The encoding method can save stored spaces and hence the cost.

Reference:

1. LAW Ka-ho(2012). *Numbers save the day*. Hong Kong: Pearson Education Asia Limited.
2. Juergen Maasz & John O'Donoghue (2011). *Real-World Problems for Secondary School Mathematics Students*. The Netherlands: Sense Publishers.
3. BBC. KS3 Computer Science - Representing text, images and sound. Retrieved from https://www.bbc.com/education/guides/zpfdwmn/revision/1

Graph Paper I

Binary number system



(i) _____

(ii) _____

(iii) _____

(iv) _____

(v) _____

(vi) _____

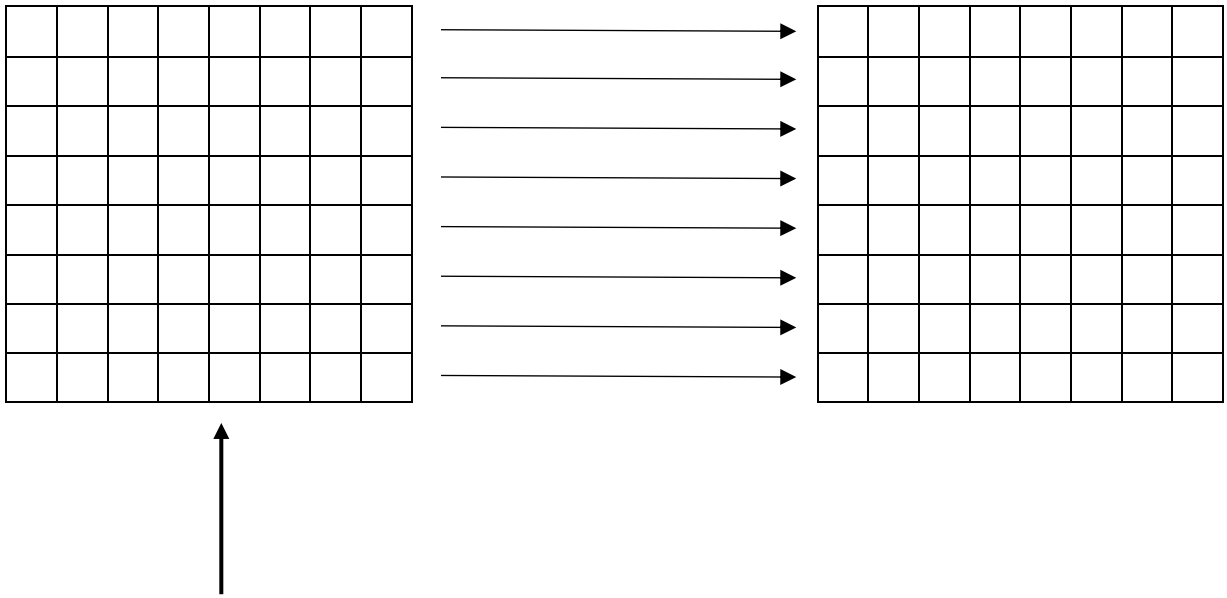
(vii) _____

(viii) _____

Denary Number system

Graph Paper II

Binary number system



(i) _____

(ii) _____

(iii) _____

(iv) _____

(v) _____

(vi) _____

(vii) _____

(viii) _____

Denary number system