INFORMATION REPRESENTATION IN COMPUTERS

Just as humans have languages with which they communicate, computers have a language that they work with; the Binary number system. Binary as a mathematical number system is a way of counting comprising of two numbers, 0 and 1. Our conventional way of counting is in tens, that is, the decimal number system, probably because we have ten fingers to represent ten numbers. In the same context, the computer is made up of switches which have only two states; ON and OFF in which it represents information.

Electricity can flow through switches: if the switch is closed, the electricity flows; if the switch is open, the electricity does not flow. To process real-world data in the computer, we need a way to represent the data in switches. Computers do this representation using a *binary coding system*. Information that can be represented in a computer include characters, integers, fixed point and floating point numbers, pictures, videos, sounds and programs.

BITS, BYTES, NIBBLES AND CHARACTERS

The word bit is a short form of **b**inary dig**it**. It's a single digit in a binary number and it can be either 1 or 0.
Bits are the building blocks for all information processing in computers.
For a sequence of two bits, we have four possible values (00, 01, 10, 11).
For a sequence of 3 bits, we have 8 possible values (000, 001, 010, 011, 100, 101, 110, 111).
This implies that for n-bits, we would have 2^n possible values.

- **A byte** is a unit of information built from bits; one byte equals 8 bits.

Bytes are combined into groups of 1 to 8 bytes called **words**.
A byte is a sequence of 8 bits, i.e., $2^8 = 256$ possible different values for a byte.
A single byte can represent many different kinds of data. What data it actually represents depends on how the computer uses the byte.

- A **nibble** is half a byte, which is a grouping of 4 bits.

- **Characters:** A character is any number, letter or symbols that a computer recognises.

Numbers such as 1, 2, 3.4, 5000…, letters or alphabets in any language or special symbols such as !, ", £, %, ^, &, * are all characters. The computer uses a byte (8 bits) to represent a single character. Every computer has a defined set of characters which they can recognise known as the character set. For a computer to be able to store and process any character, the character must be stored using a unique binary value called a character code. For example the number code for the character 'a' could be decimal 97 and a 'space' character could be 32. When a character is stored on a computer system it is therefore the number code that is actually stored as a binary number.

There are a number of standards across the globe for character codes. The code called **ASCII** (pronounced "AS-key"), which stands for **American Standard Code for Information Interchange**, uses 7 bits for each character. Since there are exactly 128 unique combinations of 7 bits, this 7-bit code can represent only characters. A more common version is ASCII-8, also called extended ASCII, which uses 8 bits per character and can represent 256 different characters. For example, the letter C is represented by the decimal value 67.

Other codes include the ANSI (American National Standards Institute), Unicode, EBCDIC etc.

THE NUMBER SYSTEMS

In the early days, humans did not have a great need to count things, they made use of words like “little” or “lots”. They used items like stones to count the number or sheep, trees or other items.

As languages developed, humans began using symbols and words for numbers. The Chinese as early as 1400BC introduced a bamboo stick decimal numbering system and developed the abacus on this system.
Later these designations were replaced in the west by the Hindu Arabic symbols for 1 through 9. But the concept of zero which the Chinese used blank space for was replaced with the symbol 0 in the Hindu-Arabic system, making it ten symbols, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9. The reason for the use of ten symbols by the Hindu-Arabic system is clearly because of humans having 10(ten) fingers on their two hands.
This explains why the Hindu-Arabic system, that is, the decimal system is the commonest numeral system. It is the modern human counting system which is based on grouping of tens and it is called base-10, the base-10 system uses ten different numeral symbols (0, 1, 2, 3, 4, 5, 6, 7, 8, 9,) to represent all numbers.
The Yuki Indians of California used a base-8 numeral system; instead of basing their system on the number of human fingers, they used the spaces between the human fingers. Also the Ancient Mayans used a base-20 system. They counted with the digits on their hands and feet because they lived in a hot climate where people didn’t wear closed toe shoes.

We can also count in numbers larger than ten. If we count in sixteen, the digits are 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F, with A = 10, B = 11, C = 12, D = 13, E = 14, F = 15. The largest digit is F. Other letters can also be used to denote numbers greater than 10 besides A, B, C, D, E, F.

The following bases are common in computing:
i) Decimal or denary (Base 10)
ii) Binary or Bicimal (Base 2)
iii) Octal (Base 8)
iv) Hexadecimal (Base 16)

A number can be converted from one base to another.

- **The Decimal Number System (Base 10)**

The decimal system is a number system to the base ten and it consists of the following digits; 0, 1, 2, 3, 4, 5, 6, 7, 8, 9. It is the commonest number system. Its importance is evident in the monetary system of countries like Nigeria and other West African countries. In Nigeria, we talk about Naira and Kobo, with ten kobo coin, forty kobo coin, one naira coin, fifty-five naira etc. The decimal system counts in **base 10.**

- **The Binary Number System (Base 2)**

The binary system has two as its base. The digits in base two are 0 and 1. Computers are built on the mechanism of a two way device system, such a system allows one to choose from two alternatives.

Examples of a two-way device systems are
a) Yes or No questions
b) True or false statements
c) Switch on or off in electrical circuit.

In such a system like on or off, of an electrical appliance, one cannot have “neither on nor off”, it must be “on or off”, in such case base two comes into play where 0 may stand for OFF and 1 may be for ON.
The binary system and other extensions of the binary system are used in computing since data is represented inside the computer in arrays of 0s and 1s only. Inside the computer, pulses of electricity represent data by combination of 0s and 1s. Most often 1 denotes circuits which are On and 0 for a circuit which is Off.
To show that a number is binary, the binary number is followed with a little 2 (as subscript).

- **The Octal Number System (Base 8)**

This number system consists of the digits 0, 1, 2, 3, 4, 5, 6, 7
We normally write the base of a number as a subscript. Example 346_8 , 1100_2 , 4150_6. Letters can also be used instead, giving 346_{eight} , 1100_{two} , 4150_{six} .
346_{eight} is read “three – four – six in base 8”.

- **Hexadecimal System (Base 16)**

A hexadecimal system uses sixteen as its base. The hexadecimal systems have the following digits 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F where A, B, C, D, E, F correspond to 10, 11, 12, 13, 14, 15.

CONVERSION FROM ONE BASE TO ANOTHER

- **Conversion from an Arbitrary Base to Base Ten**

A number system in any base can be converted to base ten by using expanded notation and Repeated/ Extended multiplication. This course will treat just the expanded notation.

Example

Convert each of the following numbers in the given base to base ten.
a) 536_8 (b) 11011_2 (c) 1243_5 (d) $B3AC_{16}$

Solution

a) 536_8
Using the expanded notation
$536_8 = 5 \times 8^2 + 3 \times 8^1 + 6 \times 8^0$
$= 5 \times 64 + 3 \times 8 + 6 \times 1$
$= 320 + 24 + 6 = 350_{ten}$

b) 11011_2
Using expanded notation
$11011_2 = 1 \times 2^4 + 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0$
$= 1 \times 16 + 1 \times 8 + 0 \times 4 + 1 \times 2 + 1 \times 1$
$= 16 + 8 + 0 + 2 + 1$
$= 27_{ten}$

Using expanded notation
c) $1243_5 = 1 \times 5^3 + 2 \times 5^2 + 4 \times 5^1 + 3 \times 5^0$
$= 1 \times 125 + 2 \times 25 + 4 \times 5 + 3 \times 1$
$= 125 + 50 + 20 + 3$
$= 198_{ten}$

Using the expanded notation

d) $B3AC_{16}$ $= 11 \times 16^3 + 3 \times 16^2 + 10 \times 16^1 + 12 \times 16^0$

$= 11 \times 4096 + 3 \times 256 + 10 \times 16 + 12 \times 1$

$= 45056 + 768 + 160 + 12 = 45996_{ten}$

- **Conversion from Base Ten to Binary (Base Two)**

To change a number from base ten to base two, one can use any of the following approaches;

i) The method of successive division

ii) Use of the correct place values in base two

In method of successive division, 2 is used to continuously divide the base ten number and the remainders collected from downwards. The division ends when the divisor (2) becomes greater than the dividend, in this case the quotient is zero and the remainder becomes the last dividend.

The other method, the use of the correct place values in base two, involves choosing the appropriate powers of 2 which when added, gives the number in base ten.
This course treats just the method of successive division.

Example: Convert the following base ten numbers into base two using the methods of successive divisions.

i) 17 (ii) 75 (iii) 324

Solution

Using the method of successive division

2	17
2	8 R 1
2	4 R 0
2	2 R 0
2	1 R 0
	0 R 1

10001_{two}

ii) Using the method of successive division

2	75
2	37 R 1
2	18 R 1
2	9 R 0
2	4 R 1
2	2 R 0
2	1 R 0
	0 R 1

1001011_{two}

iii) Using the method of successive division

$$\begin{array}{r|l} 2 & 324 \\ 2 & 162R0 \\ 2 & 81\ R\ 0 \\ 2 & 40\ R\ 1 \\ 2 & 20\ R\ 0 \\ 2 & 10\ R\ 0 \\ 2 & 5\ R\ 0 \\ 2 & 2\ R\ 1 \\ 2 & 1\ R\ 0 \\ & 0\ R\ 1 \end{array} \quad = 101000100_{two}$$

- **Conversion from Base Ten to Any base other than Base Two**

The change a whole number in base ten to any other base outside base two, the method of successive division is used.

Example:

Convert the following base ten numbers to the indicated base in bracket

i) 75 (base five) (ii) 73 (base eight)

iii) 324 (base sixteen) (iv) 96 (base six)

Solution

Using the method of successive division in each case

i)

$$\begin{array}{r|l} 5 & 73 \\ 5 & 14\ R\ 3 \\ 5 & 2\ R\ 4 \\ & 0\ R\ 2 \end{array} \quad = 243_{five}$$

ii)

$$\begin{array}{r|l} 8 & 73 \\ 8 & 9\ R\ 3 \\ 8 & 1\ R\ 1 \\ & 0\ R\ 1 \end{array} \quad = 111_{eight}$$

iii)

$$\begin{array}{r|l} 16 & 324 \\ 16 & 20R\ 4 \\ 16 & 1\ R\ 4 \\ & 0\ R\ 1 \end{array} \quad = 144_{sixteen}$$

iv)

$$\begin{array}{r|l} 6 & 96 \\ 6 & 16\ R\ 0 \\ 6 & 2\ R\ 4 \\ & 0\ R\ 2 \end{array} \quad = 240_{six}$$

- **Conversion from a non-denary base to another non-denary base**

To convert from a number which is not in base ten to another which is also not in base, one has to first convert to base ten (which serves as a link or standard base) before converting to the required base. For base two, eight and sixteen inter conversion, this may be time wasting because there is a relationship among them, so we have a shorter approach for the conversions among these bases.

Example: (a) Change 134_{five} to base two (b) Change 3245_{six} to base 9.

Solution

a) 134_{five} is first converted to base ten using the expanded notation

$134_5 = 1 \times 5^2 + 3 \times 5^1 + 4 \times 5^0$
$= 1 \times 25 + 3 \times 5 + 4 \times 1$
$= 44_{ten}$

Now 44_{ten} is converted to base 2 using successive division

2	44	
2	22 R 0	
2	11 R 0	
2	5 R 1	
2	2 R 1	
2	1 R 0	
	0 R 1	$= 101100_{two}$

$\rightarrow 134_{five} = 101100_{two}$

b) 3245_{six} is first converted to base ten

$3245_{six} = 3 \times 6^3 + 2 \times 6^2 + 4 \times 6^1 + 5 \times 6^0$
$= 3 \times 216 + 2 \times 36 + 4 \times 6 + 5 \times 1$
$= 648 + 72 + 24 + 5$
$= 739_{ten}$

Now 739_{ten} is converted to base 9

9	739	
9	82 R 1	
9	9 R 1	
9	1 R 0	$= 1011_9$
	0 R 1	

$\rightarrow 3245_{six} = 1011_9$