Data representation

This lecture describes the various ways in which computers can store and manipulate numbers and characters.

- **Bit:** The most basic unit of information in a digital computer is called a bit, which is a contraction of binary digit.
- **Byte:** In 1964, the designers of the IBM System main frame computer established a convention of using groups of 8 bits as the basic unit of addressable computer storage. They called this collection of 8 bits a byte.
- **Word:** Computer words consist of two or more **adjacent bytes** that are sometimes addressed and almost always are manipulated collectively. The word size represents the data size that is handled most efficiently by a particular architecture. Words can be 16 bits, 32 bits, 64 bits.
- **Nibbles:** Eight-bit byte can be divided into two 4-bit halves call nibbles.

مالفرق بين داته هي البيانات الاصلية(الخام)التي لم يجرى عليها اي عملية معالجه. بينما الانفرموشن هي البيانات التي تم عليها المعالجة وصارت معلومة data and information

➢ Data representation

Data representation is a major part of the software-hardware interface, Data need to be represented in a convenient way that simplifies common operations: addition, comparison, etc. and hardware implementation (cheap, fast, reliable computers)

If we want to compare, sort and organize data, we need to have a way to represent it in a computer – electronically. Even the characters you are now reading, as words, must be able to be represented and encoded electronically and

numerically. The lecture will explain how we represent and encode data in a computer.

- **Information Data Types:**

1. Binary numbers
2. Integers and Floating Point
3. Booleans (True, False)
4. Characters
5. Variables

1. Number representation

هو التمثيل الحقيقي الداخلي اي القيمة لهذا الرقم اي بمعنى الشكل يعادل المضمون

is the internal representation of numeric values in digital computer and calculator hardware and software. Normally, numeric values are stored as groupings of bits, named for the number of bits that compose them. The encoding between numerical values and bit patterns is chosen for convenience of the operation of the computer;

بشكل طبيعي قيمة الرقم تخزن على شكل مجموعة من bits وتسمى حسب الرقم المكون لها مثلا 2 تخزن بعنوان 2 لان شكل مجموع البت 010 يمثل رقم 2
كذلك مجموع البت group bits تستخدم من قبل الحاسوب لادارة ايعازات الحاسبه اي تستخدم مجموعة من الايعازات والتي تكون خاصة بعمليات الجمع و الطرح وغيرها وكلها تمثل بنظام 0,1 .

the bit format used by the computer's instruction set generally requires conversion for external use such as printing and display. Different types of processors may have different internal representations of numerical values. Different conventions are used for integer and real numbers. Most calculations are carried out with number formats that fit into a processor register, but some software systems allow representation with large numbers using multiple words of memory.

وكل هذه العمليات او المجاميع من البت تحتاج الى تحويل (محول) لتكون مفهومه من قبل الحاسوب والمستخدم مع الاستخدمات الخارجية مثل عملية الطباعة والعرض

مع اختلاف المعالجات فكل نوعه يملك نوع اوطريقة لتمثيل البيانات فيوجد معالجات تمثل البيانات بنظام 16 او الثنائي البايزري وغيرها من ذلك نستنتج ان تمثيل البيانات يعتمد على نوع المعالج

I. Binary system

Is a base-2 number system that uses two mutually exclusive states to represent information. A binary number is consist of elements called bits where each bit can be in one of the two possible states. Generally, we represent them with the numerals 1 and 0. We build binary numbers the same way we build numbers in our traditional base 10 system. However, instead of a one's column, a 10's column, a 100's column (and so on) we have a one's column, a two's columns, a four's column, an eight's column, and so on, as illustrated below.

- **Table 1. Binary**

$2^{...}$	2^6	2^5	2^4	2^3	2^2	2^1	2^0
...	64	32	16	8	4	2	1

جدول تحويل من النظام Binary الى Decimal

Ex:-

$(1011)_2 \longrightarrow (11)_{10}$

$2^3 \quad 2^2 \quad 2^1 \quad 2^0$

4 + 0 + 2 + 1 = 11

$(1100)_2 \longrightarrow (12)_{10}$

$(1001)_2 \longrightarrow (9)_{10}$

II. Decimal system

The decimal numeral system (also called base-ten and called denary) has ten as its base, which, in decimal, is written 10, as is the base in every positional numeral system. It is the numerical base most widely used by modern civilizations.

النظام العشري يتكون من اساس 10 ويبدأ 0 الى 9

$(13)_{10} \longrightarrow (1101)_2$

$(7)_{10} \longrightarrow (0111)_2$

III. Hexadecimal system

Hexadecimal refers to a base 16 number system. We use this in computer science for only one reason, it makes it easy for humans to think about binary numbers. Computers only ever deal in binary and hexadecimal is simply a shortcut for us humans trying to work with the computer. To represent 16 different patterns in binary, we would need exactly four bits. Therefore, each hexadecimal numeral represents exactly four bits. You should consider it an exercise to learn the following table off by heart.

Table 2. Hexadecimal, Binary and Decimal

Hexadecimal	Binary	Decimal
0	0000	0
1	0001	1
2	0010	2
3	0011	3
4	0100	4
5	0101	5
6	0110	6
7	0111	7
8	1000	8
9	1001	9
A	1010	10
B	1011	11
C	1100	12
D	1101	13
E	1110	14
F	1111	15

Example: - Convert $(9)_{16}$ ⟶ $(1001)_2$

$(9)_{16}$ ⟶ $(1001)_2$

Example:-

Convert $(4E)_{16}$ to binary:

$(4)_{16} = (0100)_2$

$(E)_{16} = (1110)_2$

So

$(4E)_{16} = (01001110)_2$

Example:-

Convert (4A01)16 to binary:

$(4)_{16} = (0100)_2$

$(A)_{16} = (1010)_2$

$(0)_{16} = (0000)_2$

$(1)_{16} = (0001)_2$

Example 2

$(4A01)_{16} = (0100101000000001)_2$

2. Characters Representing

Characters need to be encoded in binary too, Operations on characters have simpler requirements than on numbers, so the encoding choice is not crucial, most common representation is **ASCII** – Each character is held in a byte.

➢ What is ASCII?

ASCII, stands for the American Standard Code for Information Interchange. ASCII was originally based on the English alphabet and consists of 128 characters including A-Z, 0-9, punctuation, spaces, and other control codes that can be found on a standard English keyboard. These 128 characters are then assigned a number from 0 to 127 to represent them in data transfer from one computer to the other.

http://www.computerhope.com/jargon/a/ascii.htm

Char	Dec	Binary	Char	Dec	Binary	Char	Dec	Binary
!	033	00100001	A	065	01000001	a	097	01100001
"	034	00100010	B	066	01000010	b	098	01100010
#	035	00100011	C	067	01000011	c	099	01100011
$	036	00100100	D	068	01000100	d	100	01100100
%	037	00100101	E	069	01000101	e	101	01100101
&	038	00100110	F	070	01000110	f	102	01100110
'	039	00100111	G	071	01000111	g	103	01100111
(	040	00101000	H	072	01001000	h	104	01101000
)	041	00101001	I	073	01001001	i	105	01101001
*	042	00101010	J	074	01001010	j	106	01101010
+	043	00101011	K	075	01001011	k	107	01101011
,	044	00101100	L	076	01001100	l	108	01101100
-	045	00101101	M	077	01001101	m	109	01101101

.	046	00101110	N	078	01001110	n	110	01101110
/	047	00101111	O	079	01001111	o	111	01101111
0	048	00110000	P	080	01010000	p	112	01110000
1	049	00110001	Q	081	01010001	q	113	01110001
2	050	00110010	R	082	01010010	r	114	01110010
3	051	00110011	S	083	01010011	s	115	01110011
4	052	00110100	T	084	01010100	t	116	01110100
5	053	00110101	U	085	01010101	u	117	01110101
6	054	00110110	V	086	01010110	v	118	01110110
7	055	00110111	W	087	01010111	w	119	01110111
8	056	00111000	X	088	01011000	x	120	01111000
9	057	00111001	Y	089	01011001	y	121	01111001
:	058	00111010	Z	090	01011010	z	122	01111010
;	059	00111011	[	091	01011011	{	123	01111011
<	060	00111100	\	092	01011100	|	124	01111100
=	061	00111101	]	093	01011101	}	125	01111101
>	062	00111110	^	094	01011110	~	126	01111110
?	063	00111111	_	095	01011111	_	127	01111111
@	064	01000000	`	096	01100000			

https://www.dynadot.com/community/help/question/what-is-ascii