"A Study and Analysis of Digital Mathematical Application in Data Processing, Representation and Communication in Computers"

Suma

Faculty of Mathematics, Mysore.

1. Introduction :

The research would deal with the understanding of binary –digital mathematical applications in computers.

The research would first analyse the mathematical applications such as

1) Binary mathematics, hexadecimal, binary coded decimal maths, binary coded decimal interchange code direct maths, etc..(in arithmetic and logic unit) and
2) Mathematical application in the data representations ASCII, BCD, EBCDIC etc..
3) Mathematical application in the transmission of data –encryption and decryption in computers.

The researcher would mainly focus upon understanding of the mathematical application in computer – processing, data representation and transmission which entirely different from regular mathematics.The present research journal paper will try to brings out the strength and weakness of these applications and limitations and suggest newer methods and sequences which will be able bring better efficiency in the data representation, coding (encryption and decryption), and transmission and vital data security .

2. The objective of the present research :

The research brings the analysis of how the digital mathematics is being used for the digital electronics and the design of computers and the mathematical applications are used in the storage, processing, data representation and transmission of information in computers. The objectives of the research is to find the limitations, strength and weakness of the application.

3. Literature survey and review of literature of previous research papers in the field of research:

3.1 The use of two digits 0 and 1 is used to represent any data in computers. Because however complex is the electrical circuits, the IC chips and microprocessors which form the computers – they can only understand two states of being i.e "the presence or absence of electrical pulse", "high or low voltage", "positive movement of electricity or reverse movement" everything has to be represented .

3.2 The name "ASCII" stands for "American Standard Code for Information Interchange", which was a particular way of assigning bit patterns to the characters on a typewriter. The ASCII system even includes "characters" for ringing a bell (useful for getting attention on old telegraph systems), deleting the previous character (kind of an early "undo"), and "end of transmission" (to let the receiver know that the message was finished). These days those characters are rarely used, but the codes for them still exist (they are the missing patterns in the table above). Nowadays ASCII has been surplanted by a code called "UTF-8", which happens to be the same as ASCII if the extra left-hand bit is a 0, but opens up a huge range of characters if the left-hand bit is a 1.

There are several other codes that were popular before ASCII, including the Baudot code and EBCDIC. A widely used variant of the Baudot code was the “Murray code”, named after New Zealand born inventor Donald Murray. One of Murray’s significant improvements was to introduce the idea of “control characters”, such as the carriage return (new line). The “control” key still exists on modern keyboards.

References: http://www.csfieldguide.org.nz/releases/1.9.9/DataRepresentation.html

The article “A Tutorial on Data RepresentationIntegers, Floating-point Numbers, and Characters” explains the number syetems and evolution of its usage in compute dat representation.

Number Systems: Human beings use *decimal* (base 10) and *duodecimal* (base 12) number systems for counting and measurements (probably because we have 10 fingers and two big toes). Computers use *binary* (base 2) number system, as they are made from binary digital components (known as transistors) operating in two states - on and off. In computing, we also use *hexadecimal* (base 16) or *octal* (base 8) number systems, as a *compact* form for represent binary numbers.

Decimal (Base 10) Number System: Decimal number system has ten symbols: 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9, called *digit*s. It uses *positional notation*. That is, the least-significant digit (right-most digit) is of the order of 10^0 (units or ones), the second right-most digit is of the order of 10^1(tens), the third right-most digit is of the order of 10^2(hundreds), and so on. For example, 735 = 7×10^2 + 3×10^1 + 5×10^0.

Binary (Base 2) Number System: Binary number system has two symbols: 0 and 1, called *bits*. It is also a *positional notation*, for example,10110B = 1×2^4 + 0×2^3 + 1×2^2 + 1×2^1 + 0×2^0. It denote a binary number with a suffix B. Some programming languages denote binary numbers with prefix 0b(e.g., 0b1001000), or prefix b with the bits quoted (e.g., b'10001111').A binary digit is called a *bit*. Eight bits is called a *byte* (why 8-bit unit? Probably because 8=23).

Hexadecimal (Base 16) Number System: Hexadecimal number system uses 16 symbols: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, and F, called *hex digits*. It is a *positional notation*, for example, A3EH = 10×16^2 + 3×16^1 + 14×16^0. It denote a hexadecimal number (in short, hex) with a suffix H. Some programming languages denote hex numbers with prefix 0x (e.g., 0x1A3C5F), or prefix x with hex digit quoted (e.g., x'C3A4D98B').Each hexadecimal digit is also called a *hex digit*. Most programming languages accept lowercase 'a' to 'f' as well as uppercase 'A' to 'F'.Computers uses binary system in their internal operations, as they are built from binary digital electronic components. However, writing or reading a long sequence of binary bits is cumbersome and error-prone. Hexadecimal system is used as a *compact* form or *shorthand* for binary bits. Each hex digit is equivalent to 4 binary bits, i.e., shorthand for 4 bits, as follows:

Big Endian vs. Little Endian: Modern computers store one byte of data in each memory address or location, i.e., byte addressable memory. An 32-bit integer is, therefore, stored in 4 memory addresses. The term "Endian" refers to the *order* of storing bytes in computer memory. In "Big Endian" scheme, the most significant byte is stored first in the lowest memory address (or big in first), while "Little Endian" stores the least significant bytes in the lowest memory address. For example, the 32-bit integer 12345678H (221505317010) is stored as 12H 34H 56H 78H in big endian; and 78H 56H 34H 12H in little endian. An 16-bit integer 00H 01H is interpreted as 0001H in big endian, and 0100H as little endian.

Main Ref: Chua Hock-Chuan (ehchua@ntu.edu.sg)

References & Resources:

i. (Floating-Point Number Specification) IEEE 754 (1985), "IEEE Standard for Binary Floating-Point Arithmetic".
ii. (ASCII Specification) ISO/IEC 646 (1991) (or ITU-T T.50-1992), "Information technology - 7-bit coded character set for information interchange".
iii. (Latin-I Specification) ISO/IEC 8859-1, "Information technology - 8-bit single-byte coded graphic character sets - Part 1: Latin alphabet No. 1".
iv. (Unicode Specification) ISO/IEC 10646, "Information technology - Universal Multiple-Octet Coded Character Set (UCS)".
v. Unicode Consortium @ http://www.unicode.org.

3.3 The electronic data representation -

Rational choices in the electronic representation of data presuppose their analysis in terms of criteria such as the following:

i. nature of the data
ii. structure of the data
iii. demands on access and retrieval of the data.
iv. Kinds of data

One major distinction between kinds of data is between digital and analog data. Digital data code everything they represent in terms of figures (inside the computer, in binary digits); analog data bear an iconic resemblance to the phenomena they represent. Digital computers can only process digital data (because a bit is either set or not set, but not 37% set). Analog data therefore have to be digitalized before they enter the digital computer (and they may be reanalogized if they leave it in order to be perceived).When talking of computerized (digital) data, we will nevertheless distinguish between analog and symbolic data, according to their function for the user:

I. Analog (computer) data are digitalized analog (raw) data. They represent phenomena to be perceived (visually or auditorily). Examples include pictures, video and audio records.
II. Symbolic data are treated as signs by the user. Examples include textual data and figures; both are called alphanumeric data.

The technical counterpart to these various kinds of data are data types. These are specific constellations and interpretations of sequences of bits and bytes. Here are some examples of data types used at the beginning of the 3rd millennium:

I. analog data types: BMP, MPEG, RAW, WAV, MP3 ...
II. symbolic data types: integer, long integer, date, one-byte character, two-byte character, (character) string ...

The above analog data types are, at the same time, file types, which reflects the fact that the internal structure of such files is generally opaque to the user. Symbolic data can always be represented as text; analog data cannot. That means that the user can, within certain limits, choose the data type for his symbolic data.

Ref.www.christianlehmann.eu

3.4 : Digital data, in information theory and information systems, are discrete, discontinuous representations of information or works, as contrasted with continuous, or analog signals which behave in a continuous manner, or represent information using a continuous function.Although digital representations are the subject matter of discrete mathematics, the information represented can be either discrete, such as numbers and letters, or it can be continuous, such as sounds, images, and other measurements.

The word digital comes from the same source as the words digit and digitus (the Latinword for finger), as fingers are often used for discrete counting. Mathematician George Stibitz of Bell Telephone Laboratories used the word digital in reference to the fast electric pulses emitted by a device designed to aim and fire anti-aircraft guns in 1942.[1] The term is most commonly used in computing and electronics, especially where real-world information is converted to binary numeric form as in digital audio and digital photography.

Ref: Ceruzzi, Paul E (June 29, 2012). Computing - A Concise History. MIT Press. ISBN 978-0-262-51767-6.

3.5 review of literature of previous research papers in the field of research:

Richard C. Gossweiler, IIIstate in the research article “N-space indexing of digital data representations using physical tags “states that “A system for N-space navigation of digital data sets comprising an electronic tag having a digitally readable identifier, an electronic tag reader configured to read the identifier of the electronic tag, and a computing system connected to the electronic tag reader to provide digital navigation services of N-space data sets in response to reading the unique identifier of each electronic tag.

Field Of The Invention: The present invention relates to indexing complex data sets. More particularly, the present invention relates to use of separate physical identifier tags that store preset or user determined information to assist in N-space indexing and selection.

The Summary Of The Invention :Interactive navigation by users of digitally represented data spaces of N dimensions can be difficult due to the complexity of choices, awkward user interfaces, mismatched navigational input devices, or limitations in user memory or skill. For example, a user attempting to interactively comprehend and navigate through a multidimensional digital data set can become easily confused without predefined anchor points that mark known starting points, waypoints, or end points. Such anchor points may include simple time points (e.g. for navigation of audio data); centering on a preferred coordinate axis from a predefined viewpoint (e.g. for navigation of static two dimensional data displays); object centered movable viewpoint (e.g. for “flyby” tracking of rendered three dimensional objects); or viewing of predetermined two dimensional slices of high dimension (e.g. four or more) phase spaces.

Ref :Richard C. Gossweiler, III “N-space indexing of digital data representations using physical tags (US 6422474 B1)” publication number US6422474 B1 ,Jul 23, 2002 .Xerox Corporation.

Robert V. Dickinson and Louis M. GalieMethod “ means using digital data processing means for locating representations in a stored textual data base “ state that “The method uses digital data processing means and stored representations of a table of textual block identifiers for locating in a stored textual data base those textual blocks having the best match with a query. Textual block identifiers each provide an indication of a textual block in a stored data base which contains the corresponding word. The method comprises the following steps: A query word is received having representations of a plurality of words to be located in textual blocks in the stored data base. For each of a plurality of the query words, determine a corresponding set of equivalent words which are contained in the stored data base. Each set of equivalent words is equivalent to the corresponding query word. Each equivalent word has a corresponding group of textual block identifiers represented in the stored table. Process the representations of the textual block identifiers in those groups which correspond to the determined equivalent words to thereby form a score for at least one textual block. The score provides an indication of the total number of the sets which have at least one equivalent word in the at least one textual block. The score is utilized to provide output data pertaining to selected textual blocks in the stored textual data base.”

Ref : Robert V. Dickinson and Louis M. GalieMethod “ means using digital data processing means for locating representations in a stored textual data base “US4495566 A,Jan 22, 1985,System Development Corporation

The research article in IEEE Computer Graphics and Applications state that “The process of matching the type of information of interest with the ability of the different representation properties to convey this information is formalized in a generic framework. Visualization approaches are discussed, and the natural scene paradigm on which the methodology for choosing a representation is based is described. Data types and interpretation aims are explored. The use of the paradigm, the matching process, and implementation and display of the representation are described. The advantages of the methodology over ad hoc display approaches are examined.

Ref: research article Published in” IEEE Computer Graphics and Applications”, Volume: 11, Issue: 3, May 1991, Page(s): 56 – 67, Date of Publication: 06 August 2002, ISSN: 0272-1716. Publisher: IEEE, Sponsored by: IEEE Computer Society

3.8. Test system for acquiring, calculating and displaying representations of data sequences

Yih-ChyunJenq “Test system for acquiring, calculating and displaying representations of data sequences “states in his invention that “A system for testing electronic devices includes a waveform generator, a data acquisition system, and a computer. The waveform generator continuously generates a test signal having adjustable parameters set by the computer in response to user input. The data acquisition system acquires data representing the output of the device under test in response to the input signal and stores the last N acquired data values. The computer transfers a data sequence from the acquisition system to another memory and generates in a window on a terminal screen a wagveform display representing the stored data sequence. The computer also displays menu items referencing mathematical operations that may be performed on one or more data sequences. When a user selects one of the menu items, the computer prompts the user to select one or more windows containing waveform displays. Thereafter, the computer performs the selected operation on the data sequence controlling the waveform displays in the selected windows. When the result of the operation is a new data sequence, the computer stores the new data sequence in memory and then produces in a user-selected window a new waveform display based on the new data sequence.

Ref: Yih-ChyunJenq “Test system for acquiring, calculating and displaying representations of data sequences “ US5081592 A, invention article from Tektronix, Inc. , Jan 14, 1992

4. Research methodology :

The research methodology would include the understanding of the previous research and inventions of utility of the digital mathematical applications in computer and data representations and find the areas in which is through understanding, inventions is required . The research will focus on to find the advantages, qualities, attributes and limitations of the present usuage of digital system in computer and find better methods and representations and sequences which could be better utilized for data representation, data testing, coding, encrypting and decrypting and for the purpose of security of the vital data and information transmitted through the computers.

5. The objective of the research :

The major and overall objective of the present research is a study and analysis of digital mathematical application in Data processing, representation and communication in computers.

The broad perspective and detailed objectives of the research can be put across as follows:

1) To analyse the present mathematical applications in computer data representation (within computer , networks and internet and intranet systems)
a) In the Binary,hexadecimal, binary coded decimal, binary coded decimal interchange code direct mathematics(ASCII, BCD, EBCDIC systems) etc.
b) In protocol systems for the transmission of data.
c) In the security of the data in processing, transmission and storage.
2) The research will focus on to find the advantages, qualities, attributes and limitations of the present usage of digital mathematical system in computer for processing, storage , transmission ,coding (encryption and decryption)and security systems(sequencing codes) in standalone computer and network systems .
3) The research is to find better methods, data representations and sequences which could bring better efficiency in data representation, processing, coding (data encrypting and decrypting) and for the purpose of security of the vital data and information transmitted through the computers.

6. Conclusions :

The initial research on the digital mathematical applications in computer and data representations found that

a) There is continuous need of research to find better methods and mathematical systems of data representation for processing of data in computers and network systems.
b) The researcher has studied the complete present systems of digital mathematics used for data presentation in the computer systems.
c) The researcher found that there is great scope for improvement of the systems of sequences used to protect the data in transmission and codes used in protocol management and codes used for encryption and decryption of data used in mailing servers (goggle, yahoo etc..) whatsapp, face-book and other social networks .
d) The researcher found that there is urgent need for improvement of storage and transmission systems in data-mining and cloud computing systems used in corporate data processing and storage.
e) The researcher found that the same protocol for transmission (generally used for regular messaging and other purposes) is used while transmitting to and fro the vital banking and financial transactions , Dmat data transmissions, ATM systems and debit and credit card swiping machines leading to security lapses and cyber crimes. This area needs the computer and network system to switch on to different mode of coding (digital data representation) while these type of special data is transmitted.

The researcher strongly feel and is confident that his research would help to the betterment of the efficiency of data representation, processing, storage, transmission and security in standalone as well as in network systems.

================================