

CHAPTER-4

डाटा निरूपण

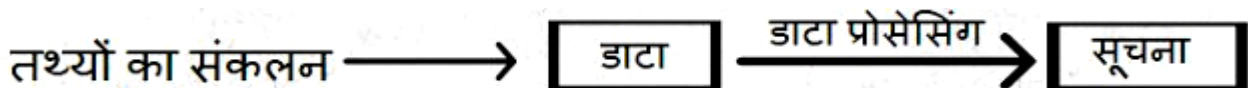
(DATA REPRESENTATION)

कंप्यूटर सूचना एवं डाटा - Information and Data

Computer की भाषा में Data सूचनाओं, तथ्य तथा आंकड़ों का संग्रह होता है कंप्यूटर द्वारा प्रोसेस की जा सकने वाली सभी प्रकार की जानकारी डाटा कहलाती है इस डाटा में किसी व्यक्ति का नाम, वस्तु का नाम, स्थान का नाम उससे जुड़े हुए विवरण और आंकड़े हो सकते हैं डाटा किसी भी रूप में हो सकता है जैसे टैक्स्ट, ऑडियो, वीडियो और फोटो

डाटा प्रोसेसिंग (Data processing)

किसी सूचना को प्राप्त करने के लिये हम डाटा (Data) यानि आंकड़े इकट्ठा करते हैं और इस डाटा (Data) का अपनी जरूरत के हिसाब से विश्लेषण या डाटा प्रोसेसिंग करते हैं इस विश्लेषण यानि **डाटा प्रोसेसिंग** की प्रक्रिया में डाटा से अर्थपूर्ण तथ्य, अंक या सांख्यिकी डाटा प्राप्त होता है इसी अर्थपूर्ण डाटा को सूचना कहते हैं जिसे उपयोग में लाया जाता है



डाटा प्रोसेसिंग (Data processing) या विश्लेषण - डाटा की उपयोगिता के आधार पर किया जाने वाला विश्लेषण डाटा प्रोसेसिंग कहते हैं डाटा से सूचना निकलने के लिए हमें बहुत सी क्रियाएं जैसे जोड़ना और घटाना करनी पड़ती है, उन सब क्रियाओं को **डाटा प्रोसेसिंग (Data processing)** कहा जाता है, इसके बाद ही सूचना प्राप्त होती है

Number system used in computer

मनुष्य गणना के लिए दशमलव आधारी संख्या पद्धति (Decimal Number System) का प्रयोग करता हैं। जिसमें 0 से 9 तक (कुल 10) अंकों का प्रयोग किया जाता हैं। अन्य सभी अंक इन्हीं अंकों से मिलकर बनते हैं। परन्तु कम्प्यूटर Decimal Number System का प्रयोग नहीं करता हैं। कम्प्यूटर में प्रयोग होने वाली Number systems हैं।

- Binary Number System
- Octal Number System
- Decimal Number System
- Hexadecimal Number System

Number System	आधार (Base)	कुल अंक (Total Number)	महत्तम अंक (Highest Digit)
Binary Number	2	0,1	1

System			
Octal Number System	8	0,1,2,3,4,5,6,7	7
Decimal Number System	10	0,1,2,3,4,5,6,7,8,9	9
Hexadecimal Number System	16	0,1,2,3,4,5,6,7,8,9, A,B,C,D,E,F	F

Binary Number System

कम्प्यूटर एक इलेक्ट्रॉनिक मशीन है जो विद्युत धारा पर कार्य करता है। यह केवल दो ही परिस्थितियों को जान सकता है। पहला, जब Circuit में धारा प्रवाहित हो रही है अर्थात् Circuit का स्विच ऑन है तो इसे संकेत '1' कहा जाता है। दूसरी स्थिति में Circuit में धारा प्रवाहित नहीं हो रही है, अर्थात् Circuit का स्विच ऑफ है तो इसे संकेत '0' कहा जाता है। इससे हम कह सकते हैं कि कम्प्यूटर केवल Binary Number System में सभी संख्याएं दो अंक 0 तथा 1 का प्रयोग कर लिखी जाती हैं। इसी कारण कम्प्यूटर को डाटा या निर्देश देने से पहले उसे 0 या 1 (ऑफ या ऑन) में बदलना पड़ता है।

(Octal Number System)

Octal number system का बेस 8 होता है क्योंकि इसमें केवल 8 अंक 0 से 7 तक होते हैं। ये 8 अंक निम्नलिखित हैं (0, 1, 2, 3, 4, 5, 6, 7) होते हैं।

Decimal Number System (डेसीमल नंबर सिस्टम)

डेसीमल नंबर सिस्टम का प्रयोग हम अपनी दैनिक जीवन में करते हैं, जिसमें किसी भी संख्या को प्रदर्शित करने के लिए 0, 1, 2, 3, 4, 5, 6, 7, 8, तथा 9 अंक प्रयोग किये जाते हैं। इस Number System का आधार (base or radix) 10 है। बेस किसी नंबर सिस्टम में प्रयोग किये जाने वाले अंकों की संख्या होती है, डेसीमल नंबर सिस्टम में दस अंक होते हैं इसी कारण इसका बेस भी 10 होता है।

Hexadecimal Number System (हेक्साडेसीमल नंबर सिस्टम)

हेक्साडेसिमल संख्या पद्धति में 0 से 15 अर्थात् 16 अंकों का उपयोग किया जाता है। इनमें 0 से 9 अंक तथा 10 से 15 को A to F अल्फाबेट से निरूपित किया जाता है। हेक्साडेसिमल संख्या को बाइनरी के चार अंकों के समूह से दर्शाया जाता है, जैसे – हेक्साडेसिमल 1 संख्या का बाइनरी मान 0001 होता है। इसमें आधार 16 होता है इस नंबर सिस्टम में कुल 16, Alphanumeric वैल्यू होती है इसलिए इसका बेस 16 होता है।

कम्प्यूटर कोड (Computer Codes)

VIMS EDUCATION INSTITUTE

कम्प्यूटर में डाटा अक्षरों (Alphabets), विशेष चिन्हों (Special Characters) तथा अंकों (Numeric) में हो सकता है। अतः इन्हें अल्फान्युमेरिक डाटा (Alphanumeric Data) कहा जाता है। डाटा में प्रत्येक अक्षर, चिन्ह या अंक को एक विशेष कोड द्वारा व्यक्त किया जाता है।

बाइनरी कोडेड डेसिमल (BCD – Binary Coded Decimal)

इसमें संपूर्ण Decimal Number को Binary में बदलने की बजाय Decimal Number के प्रत्येक अंक को उसके चार अंकीय बाइनरी तुल्यांक(Rating)से प्रतिस्थापित (Substituted)कर दिया जाता है। इसे 4 Bit BCD Code कहा जाता है।

आस्की (ASCII – American Standard Code for Information Interchange)

आस्की (ASCII) एक लोकप्रिय कोडिंग सिस्टम है जिसका प्रारंभ आन्सी (ANSI – American National Standards Institute) द्वारा 1963 में किया गया। इसमें एक कैरेक्टर के लिए 8 बिट और तीव्र निरूपण के लिए Hexadecimal number System का प्रयोग किया गया। कम्प्यूटर के की-बोर्ड में प्रयुक्त प्रत्येक कैरेक्टर के लिए एक विशेष आस्की कोड निर्धारित किया गया है। इसमें एक कैरेक्टर के लिए 8 बिट का प्रयोग किया जाता है।

यूनीकोड (Unicode-Universal Code)

कम्प्यूटर के बढ़ते व्यवहार तथा अलग-अलग भाषाओं में कम्प्यूटर के उपयोग ने एक Public code की आवश्यकता को जन्म दिया जिसमें संसार के प्रत्येक कैरेक्टर के लिए एक अलग कोड निर्धारित हो ताकि प्रत्येक भाषा, प्रत्येक प्रोग्राम तथा प्रत्येक साफ्टवेयर में उसका प्रयोग किया जा सके। इसके लिए यूनीकोड की व्यवस्था की गई जिसमें एक लाख कैरेक्टर के निरूपण की क्षमता है।

यूनीकोड विश्व की सभी भाषाओं में प्रयुक्त पहले 256 कैरेक्टर का निरूपण आस्की कोड के समान ही है। इसमें प्रत्येक कैरेक्टर को 32 बिट में निरूपित किया जाता है। यूनीकोड में तीन प्रकार की व्यवस्था प्रयोग में लायी जाती हैं।

i. यूटीएफ – 8 (UTF-8-Unicode Transformation Format-8)

यूटीएफ-8 फार्मेट में समस्त यूनीकोड अक्षरों को एक, दो, तीन या चार बाइट के कोड में बदला जाता है।

ii. यूटीएफ – 16 (UTF-16)

इस फार्मेट में यूनीकोड अक्षरों को एक या दो शब्दों (1 शब्द = 16 बिट) के कोड में बदला जाता है। अतः इसे Word Oriented Format भी कहते हैं।

iii. यूटीएफ-32 (UTF-32)

इस कोड में समस्त अक्षरों को दो शब्दों (Words) यानी 32 बिट के यूनीकोड में बदला जाता है।

दशमलव	बाइनरी	ऑक्टल	हेक्सा-डेसीमल
0	0	0	0

1	1	1	1
2	10	2	2
3	11	3	3
4	100	4	4
5	101	5	5
6	110	6	6
7	111	7	7
8	1000	10	8
9	1001	11	9
10	1010	12	A
11	1011	13	B
12	1100	14	C
13	1101	15	D
14	1110	16	E
15	1111	17	F