

SECTION 4

DATA REPRESENTATION IN COMPUTERS

Data Representation

Data representation in computers refers to how information (like text, numbers, images, sound, and more) is stored and processed in a computer system. Since computers can only understand binary (1s and 0s), all types of data, regardless of their original format, must be converted into binary for the computer to process.

Encoding and Decoding

In computers, **encoding** is the process of converting data (like text, images, or sound) into binary for efficient storage or transmission. **Decoding** is the reverse process where the encoded data is converted back to its original form.

Data can be represented in computers in different ways.

Numeric Data Representation

Integers

Whole numbers (e.g., 5, -42) are represented in binary as sequences of 0s and 1s. For negative numbers, computers use representations like **two's complement** to handle the **sign** (positive or negative).

Floating-Point Numbers

Floating point numbers (float and double) are used to represent **fractional numbers** (e.g., 3.14, -0.01). Floating-point numbers are represented using a sign, exponent, and fractional part. This is similar to **scientific notation** in math, where numbers are expressed as a **coefficient** multiplied by a power of 10.

Text Representation

Character Encoding

Character encoding is the process of converting characters (like letters, digits, symbols, and punctuation) into binary.

Common Character Encodings:

- ASCII (American Standard Code for Information Interchange): Represents characters using 7 or 8 bits (e.g., 'A' = 65 in ASCII).
- Unicode: An extended system that can represent a vast range of characters from different languages, symbols, and emojis, often using 16 or 32 bits per character.

Images

Images are represented by a grid of **pixels**, with each pixel having a color value. Colors are typically represented in RGB (Red, Green, Blue) format, where each color component is assigned a value between 0 and 255.

Audio

Audio is typically represented as a sequence of samples, which are measurements of sound **amplitude** at specific **intervals**. These samples are encoded in digital formats (e.g., MP3) using a combination of binary values to represent sound waves.

Video

Video combines images (**frames**) and sound. Each frame is essentially an image, and the sound is represented as audio. Together, they are encoded in video formats (e.g., MP4, AVI), where each frame is **compressed** using encoding algorithms like H.264. Video players **decompress** these frames and **display** them sequentially while playing the video.

Data Compression

Data is often compressed to reduce **storage space** and improve transmission **efficiency**. Compression techniques can be **lossless** (e.g., ZIP, PNG), where no data is lost, or **lossy** (e.g., JPEG, MP3), where some data is discarded for smaller file sizes.