Data Storage

Data Storage

1.1 Bits and Their Storage

1.2 Main Memory

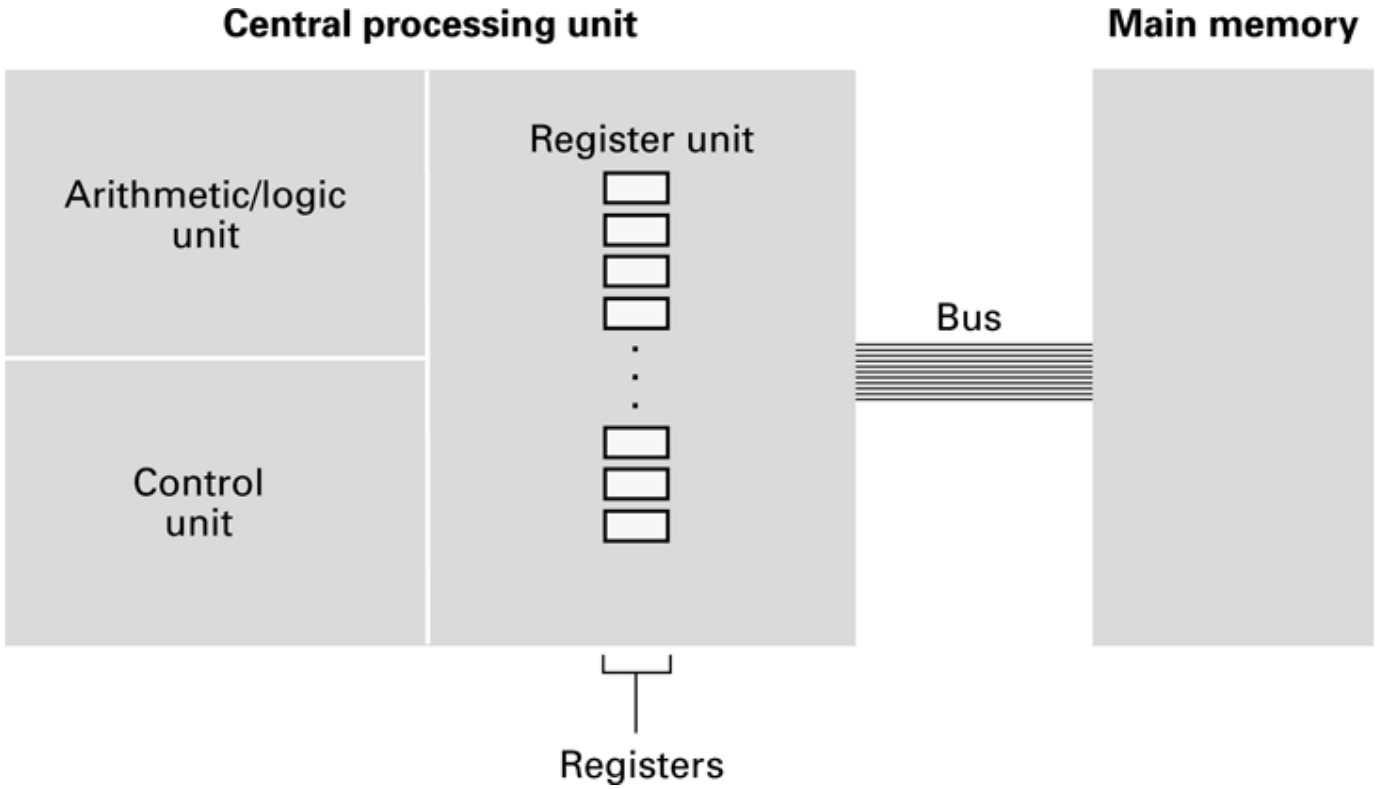
1.3 Mass Storage

1.1 Bits and Their Storage

- **The CPU is composed of two parts:**
- Control unit.
- Arithmetic Logic Unit. (ALU).

- **The control unit contains a circuitry that directs and controls to other parts of the computer.**
- **ALU contains circuitry that executes all arithmetic (+ , - , ÷ , x) and logical operations: AND, OR, XOR, NOT**
- **REGISTERS:**
  - **They are temporary storage, areas for instructions or data or address**.
  - They exist in CPU.
  - They are faster than memory.
  - They are used to:
    - Store data.
    - Accept data.
    - Transfer data.

Figure 2.1 CPU and main memory connected via a bus



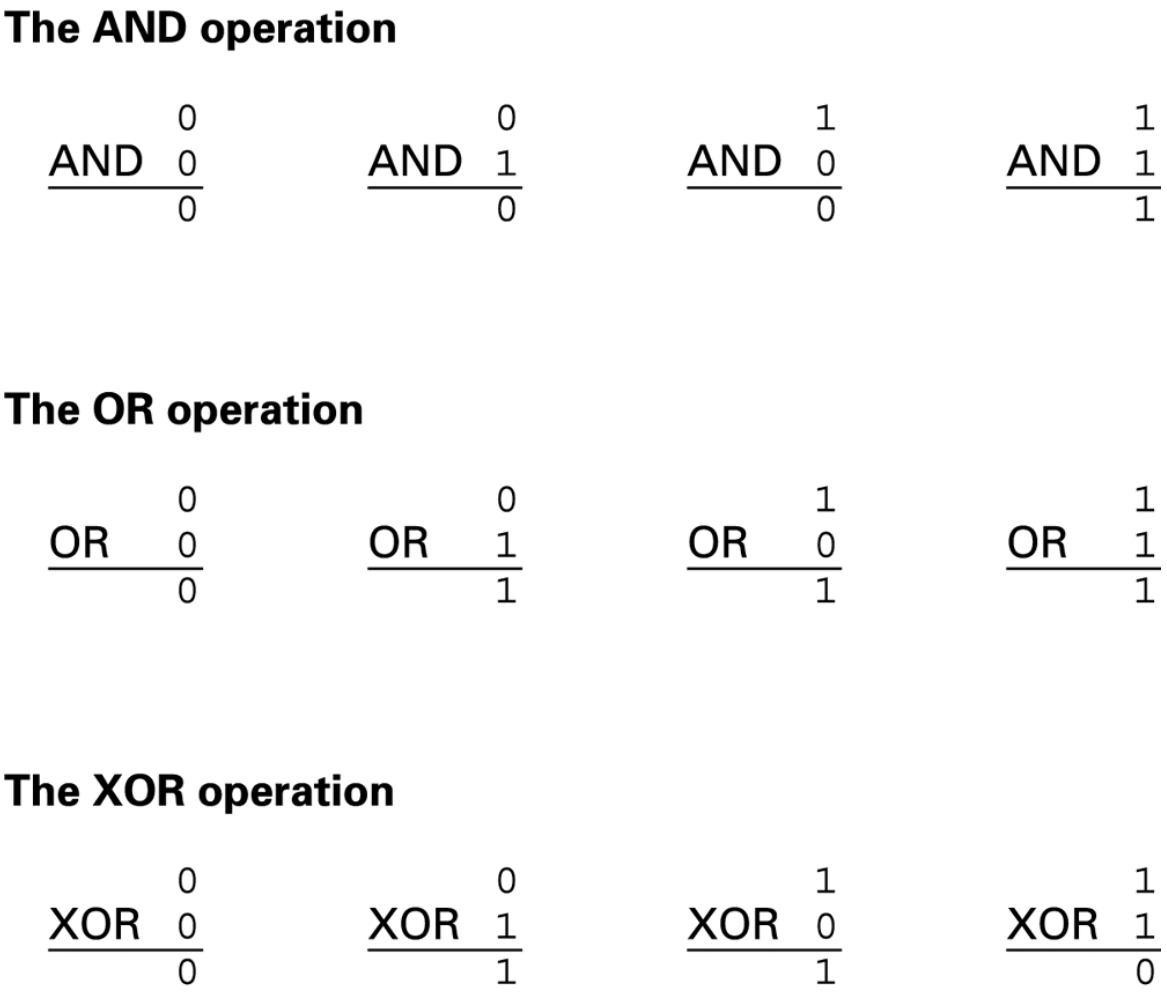
Bits and Bit Patterns

- **Bit:** Binary Digit (0 or 1)
- Bit Patterns are used to represent information.
  - Numbers
  - Text characters
  - Images
  - Sound
  - And others

Boolean Operations

- **Boolean Operation:** An operation that manipulates one or more true/false values
- Specific operations
  - AND
  - OR
  - XOR (exclusive or)
  - NOT

Figure 1.1 The Boolean operations AND, OR, and XOR (exclusive or)



Gates

- **Gate:** A device that computes a Boolean operation
  - Often implemented as (small) electronic circuits
  - Provide the building blocks from which computers are constructed
  - VLSI (Very Large Scale Integration)

As result (for an operation):

- Registers hold data immediately related to the operation being executed.
- Memory is used to store data and programs required in the near future.
- Auxiliary memory is used to store data and programs required later.

binary notation

- binary notation is a means of representing numeric values using only the digits 0 and 1 rather than the ten digits 0 through 9.

1.2 Main Memory

- **Memory:**
  - It's called also:
- * Primary memory
- * Main storage
- * Internal storage
- * Primary storage
- * Main memory
- * Internal memory
  - It holds instructions and data of the program being executed .
  - It cannot hold the program if it is not executed because:
  - Most memories destroy its data if the computer is turned off.
  - If the computer is shared, other users will need the memory space for their programs.
  - There may be no enough space to hold your program.

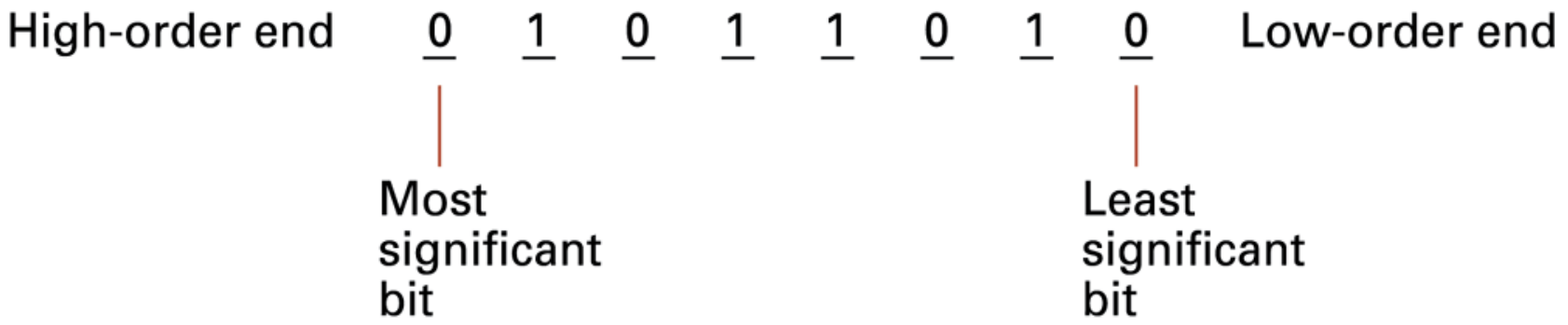
How the CU finds instructions and Data:

- There is an address for each location in memory (sometimes called numerical designator), CPU will refer to this address to store / retrieve data.
- In computer programs, symbolic addresses are used

Main Memory Cells

- **Cell:** A unit of main memory (typically 8 bits which is one **byte**)
  - **Most significant bit:** the bit at the left (high-order) end of the conceptual row of bits in a memory cell
  - **Least significant bit:** the bit at the right (low-order) end of the conceptual row of bits in a memory cell

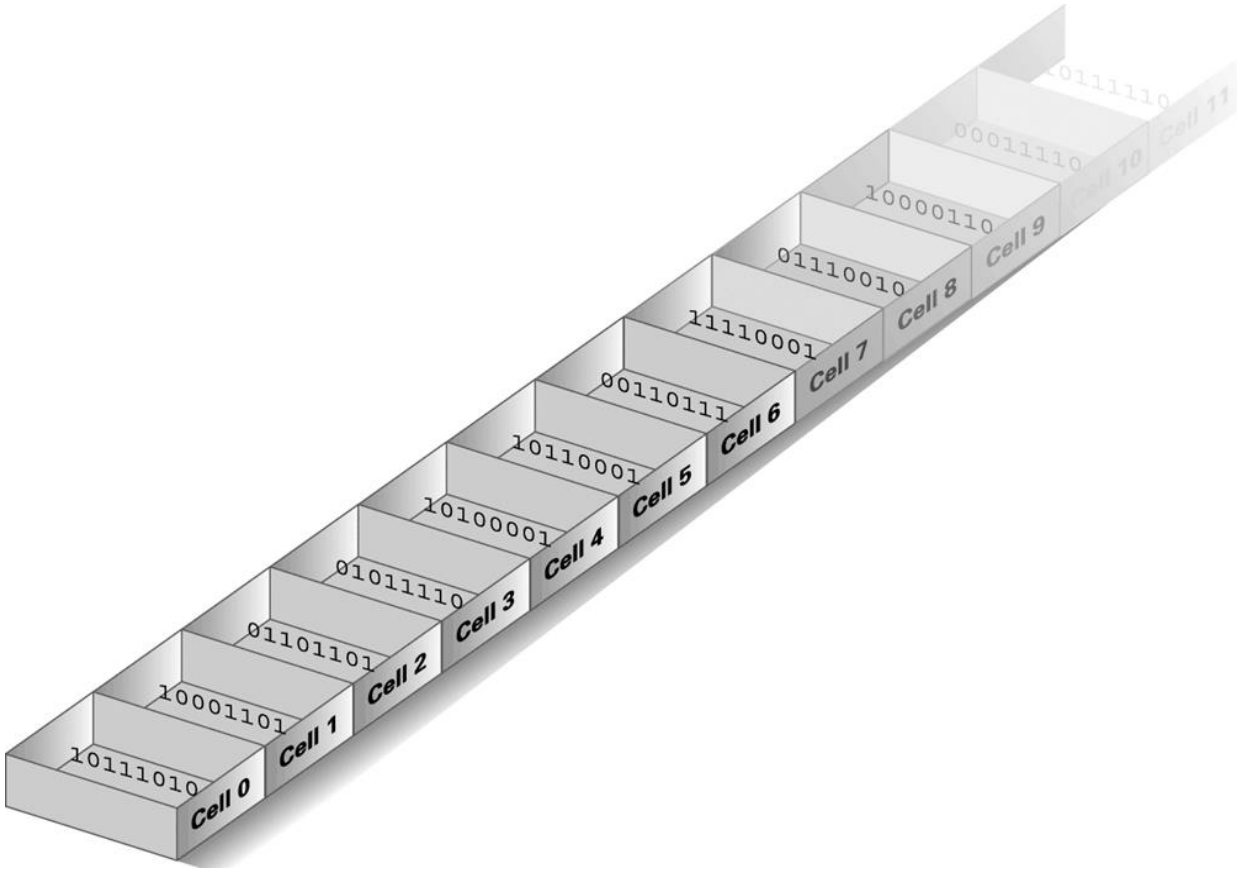
Figure 1.7 The organization of a byte-size memory cell



Main Memory Addresses

- **Address:** A "name" that uniquely identifies one cell in the computer's main memory
  - The names are actually numbers.
  - These numbers are assigned consecutively starting at zero.
  - Numbering the cells in this manner associates an order with the memory cells.

Figure 1.8 Memory cells arranged by address



Bits, Bytes and words:

- 0 or 1 is called a bit (Binary digit)
- A group of bits is called a byte. A byte = 8 bits
- 2^{10} = 1024 bytes = 1 kilobyte (KB)
- 2^{10} x 2^{10} bytes = million bytes = 1 megabytes. (MB)
- 1 billions of bytes = 1 GigaByte (GB)
- 1 trillion of bytes = 1 TeraByte (TB)
- 1 quadrillion of bytes = 1 Petabyte (PB)

RAM and ROM:

- There are two types of memory chips:
  - RAM (Random Access Memory): two types:
- ***SRAM:***
  - Static
  - Faster
  - No intervention from the CPU as long as power is maintained

- DRAM
  - Dynamic
  - Must be constantly refreshed (recharged) by the CPU
  - Used for most PC memories because its size and cost advantages

- ROM (Read – only – Memory) data cannot be changed.

1.3 Mass Storage

- **Mass (Secondary) Storage**
  - Magnetic disk storage:
  - Floppy Disks
  - Hard Disks
  - Tape
- Optical disk storage
  - Compact disks
  - DVD-ROM
  - Blue-ray Disks
- Flash Drives
  - Flash Drives
  - Secure Digital (SD) Memory Card

Secondary Storage

- Magnetic disk storage:
- ~~Floppy Disks ;~~ ***not in use now***-a-days
  - ✓ Made of flexible Mylar and coated with iron oxide
  - ✓ Magnetized spots on tracks on its surface
  - ✓ Composed of tracks and sectors
  - ✓ Popular for PCs
  - ✓ 3.5-inch with 1.44MB
  - ✓ Portable
  - ✓ Convenient for backup
  - ✓ High capacity drives: up to 750MB

1.3 Mass Storage

Secondary Storage

- Magnetic disk storage:
- Hard Disks
  - ✓ Rigid platter coated with magnetic oxide that can be magnetized
  - ✓ Varity of sizes
  - ✓ Several platters can be assembled into a disk pack
  - ✓ A Disk drive is a device that enables data to be read from/written to a disk through a read/write head on the end of an access arm
  - ✓ The read/write head does not match the surface, otherwise data are destroyed (called head crash)

Secondary Storage

- Magnetic disk storage:
- Hard Disks (HD)
  - ✓ Removable HD are available
  - ✓ Up to 500s of GB
  - ✓ Fast
  - ✓ Composed of sectors and tracks
  - ✓ A cluster is a fixed number of adjacent sectors that are treated as one unit by OS.
  - ✓ A cylinder is track n on all platters
  - ✓ To reduce access time; Related data are stored on the same cylinder *(access arm movements are reduced)*

1.3 Mass Storage

How Data Is Organized in Hard Disks (HD)

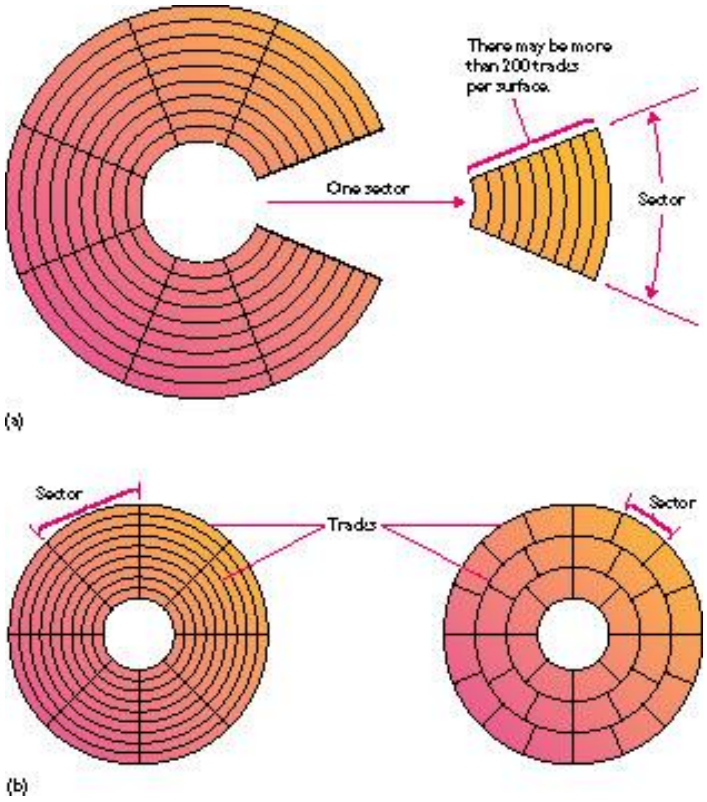
- Track
- Sector
- Cluster
- Cylinder

Track

- The circular portion of the disk surface that passes under the read/write head

Sector

- Each track is divided into small arcs called sectors on which information is recorded as a continuous string of bits
- Each track is divided into sectors that hold a fixed number of bytes
  - Typically 512 bytes per sector
- Zone recording assigns more sectors to tracks in outer zones than those in inner zones
  - Uses storage space more fully



Return

Cluster

- A fixed number of adjacent sectors that are treated as a unit of storage
  - Typically two to eight sectors, depending on the operating system

Return

Cylinder

- The track on each surface that is beneath the read/write head at a given position of the read/write heads
  - When file is larger than the capacity of a single track, operating system will store it in tracks within the same cylinder

Return

Secondary Storage

- Magnetic disk storage:
- Disk Access Speed

- **Access time=seek time + head switching + rotational delay**

- **Where:**
  - ✓ **Seek time**: is the time it takes the access arm to get into position over a particular track
  - ✓ **Head switching**: is the activation of a particular head over a particular track on a particular surface
  - ✓ **Rotational delay**: is the time for the desired data to rotate under the head.
- Data transfer: after data are found, they are transferred from disk to memory
- Disk cache can be used to improve the performance

1.3 Mass Storage

Magnetic Tape Storage

- Data is stored as extremely small magnetic spots
- 3 forms:
  - ✓ 3.5 inch tape wound on a reel
  - ✓ 3.5 inch tape in data cartridge
  - ✓ Cassette tapes
- Tape capacity: characters per inch (CPI) or Bytes per inch (BPI)
- Two heads: r/w head and erase-head

1.3 Mass Storage

Disks vs. Magnetic Tapes

- Disks are reliable
- Data on disks can be accessed directly, but tapes are sequential
- Tapes are inexpensive
- Tapes are used as a backup for data on disks

1.3 Mass Storage

Secondary Storage

- Optical disk storage:
- Metallic material spread over the surface of a disk
- Laser hits this surface to form spots that represent 0s and 1s

Compact Disks

- CD-ROM - drive can only read data from CDs
  - CD-ROM stores up to 700 MB per disk
  - Primary medium for software distribution
- CD-R - drive can write to disk once
  - Disk can be read by CD-ROM or CD-R drive
- CD-RW - drive can erase and record over data multiple times
  - Some compatibility problems trying to read CD-RW disks on CD-ROM drives

(b)

Digital Versatile Disk (DVD)

- Short wavelength laser can read densely packed spots
  - DVD drive can read CD-ROMs
  - Capacity up to 17GB
  - Allows for full-length movies
  - Sound is better than on audio CDs
- Several versions of writable and rewritable DVDs exist

1.3 Mass Storage

Secondary Storage

- Optical disk storage (*Categories*):
- Blue-ray Disks
- 5 times as the capacity of DVDs

Flash Memory

- Nonvolatile RAM
- Flash chips are used in cellular phones, digital cameras, ..
- Requires less power and smaller that disk drives
- Requires Flash Drives
- Secure Digital (SD) Memory Card

Files

- **File:** A unit of data stored in mass storage system
  - **Fields** and **keyfields**
- **Physical record** : A block of data conforming to the specific characteristics of a storage
- **Logical record**: file containing a text document would consist of paragraphs or pages. These naturally occurring blocks of data called **Logical record** (natural divisions determined by the information represented)

1.3 Mass Storage

- **File Storage and Retrieval**
- Data:
  - ✓ A Character
  - ✓ A Field: a group of characters
  - ✓ A Record: a group of fields
  - ✓ A File: a group of records
  - ✓ A Database: a group of files

1.3 Mass Storage

- **File Storage and Retrieval**
- Key: an identifying record(s)
- Buffer: a storage area used to hold data on a temporary basis, usually during the process of being transferred from one device to another.

1.4 Representation of information as bit patters

- Representing text
- Representing Numeric Values
- Representing Images
- Representing Sound

Representing Text

- **Each character (letter, punctuation, etc.) is assigned a unique bit pattern.**
  - **ASCII:** Uses patterns of 7-bits to represent most symbols used in written English text

extensions to ASCII, each of which were designed to accommodate a major language group. For example, one standard provides the symbols needed to express the text of most Western European languages. Included in its 128 additional patterns are symbols for the British pound and the German vowels a, o, and u. is simply insufficient to accommodate the alphabet of many Asian and some Eastern European languages.

- **Unicode**:. Uses patterns of 16-bits to represent the major symbols used in languages world wide. Enough to allow text written in such languages as Chinese, Japanese, and Hebrew to be represented.

Figure 1.13 The message “Hello.” in ASCII

01001000	01100101	01101100	01101100	01101111	00101110
H	**e**	**l**	**l**	**o**	**.**

Representing Numeric Values

- Binary notation: Uses bits to represent a number in base two
- Limitations of computer representations of numeric values
  - Overflow: occurs when a value is too big to be represented
  - Truncation: occurs when a value cannot be represented accurately

Representing Images

- Bit map techniques
  - Pixel: short for "picture element".
- In the case of a simple black and white image, each pixel can be represented by a single bit whose value depends on whether the corresponding pixel is black or white.
- More elaborate back and white photographs, each pixel can be represented by a collection of bits (usually eight), which allows a variety of shades of grayness to be represented.
- RGB: One byte is normally used to represent the intensity of each color component. In turn, three bytes of storage are required to represent a single pixel in the
- Analytic geometry techniques
  - Scalable
  - TrueType and PostScript

Representing Sound

- Sampling techniques

Sampling the amplitude of the sound wave at regular intervals and record the series of values obtained. using a sample rate of 8000 samples per second. Then These numeric values are then transmitted over the communication line to the receiving end.

- To obtain the quality sound reproduction obtained by today's musical CDs, a sample rate of 44,100 samples per second is used. The data obtained from each sample are represented in 16 bits (32 bits for stereo recordings).

Figure 1.14 The sound wave represented by the sequence 0, 1.5, 2.0, 1.5, 2.0, 3.0, 4.0, 3.0, 0