Chapter 1: Computer Systems

We first need to explore the fundamentals of computer processing.

Chapter 1 focuses on:
- components of a computer
- how those components interact
- how computers store and manipulate information
- computer networks
- the Internet and the World Wide Web
- programming and programming languages

Hardware and Software

- Hardware
  • the physical, tangible parts of a computer
  • keyboard, monitor, disks, wires, chips, etc.
- Software
  • programs and data
  • a program is a series of instructions
- A computer requires both hardware and software
- Each is essentially useless without the other

CPU and Main Memory

Central Processing Unit
Main Memory

Chip that executes program commands
- Intel Pentium 4 or Sun UltraSPARC III Processor

Primary storage area for programs and data that are in active use
Synonymous with RAM

Secondary Memory Devices

- Secondary memory devices provide long-term storage
- Hard disks
- Floppy disks
- 3.5" disks
- Wal lable CDs
- Tapes

Information is moved between main memory and secondary memory as needed

Input / Output Devices

- I/O devices facilitate user interaction
  - Monitor screen
  - Keyboard
  - Mouse
  - Joystick
  - Bar code scanner
  - Touch screen

- Monitor
- Keyboard
- Mouse
- Joystick
- Bar code scanner
- Touch screen
- Hard Disk
- Floppy Disk
Software Categories
- Operating System
  - controls all machine activities
  - provides the user interface to the computer
  - manages resources such as the CPU and memory
  - Windows XP, Windows 2000, Unix, Linux, Mac OS
- Application program
  - generic term for any other kind of software
  - word processors, missile control systems, games
- Most operating systems and application programs have a graphical user interface (GUI)

Analog vs. Digital
- There are two basic ways to store and manage data:
  - Analog
    - continuous, in direct proportion to the data represented
    - music on a record album: a needle rides on ridges in the grooves that are directly proportional to the voltages sent to the speaker
  - Digital
    - the information is broken down into pieces, and each piece is represented separately
    - music on a compact disc: the disc stores numbers representing specific voltage levels sampled at specific times

Digital Information
- Computers store all information digitally:
  - numbers
  - text
  - graphics and images
  - video
  - audio
  - program instructions
- In some way, all information is digitized - broken down into pieces and represented as numbers

Representing Text Digitally
- For example, every character is stored as a number, including spaces, digits, and punctuation
- Corresponding upper and lower case letters are separate characters

| Hi, Heather. |
| 72 105 44 32 72 101 97 116 104 101 114 46 |

Binary Numbers
- Once information is digitized, it is represented and stored in memory using the binary number system
- A single binary digit (0 or 1) is called a bit
- Devices that store and move information are cheaper and more reliable if they have to represent only two states
- A single bit can represent two possible states, like a light bulb that is either on (1) or off (0)
- Permutations of bits are used to store values

Bit Permutations

<table>
<thead>
<tr>
<th>1 Bit</th>
<th>2 Bits</th>
<th>3 Bits</th>
<th>4 Bits</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>00</td>
<td>000</td>
<td>0000</td>
</tr>
<tr>
<td>1</td>
<td>01</td>
<td>001</td>
<td>0010</td>
</tr>
<tr>
<td>10</td>
<td>010</td>
<td>0011</td>
<td>00110</td>
</tr>
<tr>
<td>11</td>
<td>011</td>
<td>00110</td>
<td>00111</td>
</tr>
<tr>
<td>100</td>
<td>0100</td>
<td>01001</td>
<td>0111</td>
</tr>
<tr>
<td>101</td>
<td>0101</td>
<td>01101</td>
<td>1101</td>
</tr>
<tr>
<td>110</td>
<td>0110</td>
<td>11011</td>
<td>1110</td>
</tr>
<tr>
<td>111</td>
<td>0111</td>
<td>11111</td>
<td>1111</td>
</tr>
</tbody>
</table>

Each additional bit doubles the number of possible permutations
Bit Permutations

- Each permutation can represent a particular item.
- There are $2^N$ permutations of $N$ bits.
- Therefore, $N$ bits are needed to represent $2^N$ unique items.

<table>
<thead>
<tr>
<th>Number of Bits</th>
<th>Number of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 bit</td>
<td>2^1 = 2 items</td>
</tr>
<tr>
<td>2 bits</td>
<td>2^2 = 4 items</td>
</tr>
<tr>
<td>3 bits</td>
<td>2^3 = 8 items</td>
</tr>
<tr>
<td>4 bits</td>
<td>2^4 = 16 items</td>
</tr>
<tr>
<td>5 bits</td>
<td>2^5 = 32 items</td>
</tr>
</tbody>
</table>

A Computer Specification

- Consider the following specification for a personal computer:
  - 950 MHz Pentium 4 Processor
  - 512 MB RAM
  - 30 GB Hard Disk
  - CD RW/24x / 10x / 40x
  - 17" Video Display with 1280 x 1024 resolution
  - 56 KHz Modem
- What does it all mean?

Memory

- Main memory is divided into many memory locations (or cells).
- Each memory cell has a numeric address, which uniquely identifies it.

Storing Information

- Each memory cell stores a set number of bits (usually 8 bits, or one byte).
- Large values are stored in consecutive memory locations.

Storage Capacity

- Every memory device has a storage capacity, indicating the number of bytes it can hold.
- Capacities are expressed in various units:

<table>
<thead>
<tr>
<th>Unit</th>
<th>Symbol</th>
<th>Number of Bytes</th>
</tr>
</thead>
<tbody>
<tr>
<td>kilobyte</td>
<td>KB</td>
<td>2^10 = 1024</td>
</tr>
<tr>
<td>megabyte</td>
<td>MB</td>
<td>2^20 (over 1 million)</td>
</tr>
<tr>
<td>gigabyte</td>
<td>GB</td>
<td>2^30 (over 1 billion)</td>
</tr>
<tr>
<td>terabyte</td>
<td>TB</td>
<td>2^40 (over 1 trillion)</td>
</tr>
</tbody>
</table>
RAM vs. ROM

- **RAM** - Random Access Memory (direct access)
- **ROM** - Read-Only Memory
- The term RAM and main memory are basically interchangeable
- ROM could be a set of memory chips, or a separate device, such as a CD-ROM
- Both RAM and ROM are random (direct) access devices!
- RAM probably should be called Read-Write Memory

Compact Discs

- A CD-ROM is portable read-only memory
- A microscopic pit on a CD represents a binary 1 and a smooth area represents a binary 0
- A low-intensity laser reflects strongly from a smooth area and weakly from a pit
- A CD-Recordable (CD-R) drive can be used to write information to a CD once
- A CD-Rewritable (CD-RW) can be erased and reused
- The speed of a CD drive describes how fast it can write information to a CD-R (24x), a CD-RW (10x), and how fast it can read (40x)

DVDs

- A DVD is the same size as a CD, but can store much more information
- The format of a DVD stores more bits per square inch
- A CD can store 650 MB, while a standard DVD can store 4.7 GB
  - A double-sided DVD can store 9.4 GB
  - Other advanced techniques can bring the capacity up to 17.0 GB
- There are various recordable DVD technologies - the market will determine which will dominate

The Central Processing Unit

- The CPU is on a chip called a microprocessor
- It continuously follows the fetch-decode-execute cycle:
  - Determine what the instruction is
  - Carry out the instruction
  - Retrieve an instruction from main memory
  - The speed is measured in megahertz (MHz)

The Central Processing Unit

- The CPU contains:
  - Arithmetic / Logic Unit
  - Control Unit
  - Registers
  - Small storage areas

- Performs calculations and makes decisions

- Coordinates processing steps

- The speed of a CPU is controlled by the system clock
- The system clock generates an electronic pulse at regular intervals
- The pulses coordinate the activities of the CPU
- The speed is measured in megahertz (MHz)
Monitor

- The size of a monitor (17") is measured diagonally, like a television screen.
- Most monitors these days have multimedia capabilities: text, graphics, video, etc.
- A monitor has a certain maximum resolution, indicating the number of picture elements, called pixels, that it can display (such as 1280 by 1024).
- High resolution (more pixels) produces sharper pictures.

Modem

- Data transfer devices allow information to be sent and received between computers.
- Many computers include a modulator-demodulator or modem, which allows information to be moved across a telephone line.
- A data transfer device has a maximum data transfer rate.
- A modem, for instance, may have a data transfer rate of 56,000 bits per second (bps).

Networks

- A network is two or more computers that are connected so that data and resources can be shared.
- Most computers are connected to some kind of network.
- Each computer has its own network address, which uniquely identifies it among the others.
- A file server is a network computer dedicated to storing programs and data that are shared among network users.

Network Connections

- Each computer in a network could be directly connected to every other computer in the network.
- These are called point-to-point connections.
- Adding a computer requires a new communication line for each computer already in the network.
- This technique is not practical for more than a few close machines.

Network Connections

- Most networks share a single communication line.
- Adding a new computer to the network is relatively easy.
- Network traffic must take turns using the line, which introduces delays.
- Often information is broken down into parts, called packets, which are sent to the receiving machine and then reassembled.

Local-Area Networks

- A Local-Area Network (LAN) covers a small distance and a small number of computers.
- A LAN often connects the machines in a single room or building.
Wide-Area Networks

A Wide-Area Network (WAN) connects two or more LANs, often over long distances. A LAN usually is owned by one organization, but a WAN often connects groups in different countries.

The Internet

The Internet is a WAN which spans the entire planet. The word Internet comes from the term internetworking, which implies communication among networks. It started as a United States government project, sponsored by the Advanced Research Projects Agency (ARPA) - originally it was called the ARPANET. The Internet grew quickly throughout the 1980s and 90s. Less than 600 computers were connected to the Internet in 1983; by the year 2000 there were over 10 million.

TCP/IP

A protocol is a set of rules that determine how things communicate with each other. The software which manages Internet communication follows a suite of protocols called TCP/IP. The Internet Protocol (IP) determines the format of the information as it is transferred. The Transmission Control Protocol (TCP) dictates how messages are reassembled and handles lost information.

IP and Internet Addresses

Each computer on the Internet has a unique IP address, such as: 139.179.21.213. Most computers also have a unique Internet name, which also is referred to as an Internet address: gordion.cs.bilkent.edu.tr. The first part indicates a particular computer (gordion). The rest is the domain name, indicating the organization (cs.bilkent.edu.tr).

Domain Names

The last part of each domain name, called a top-level domain (TLD) indicates the type of organization:
- edu - educational institution
- com - commercial entity
- org - non-profit organization
- net - network-based organization

Sometimes the suffix indicates the country: New TLDs have recently been added.
- tr - Turkey
- uk - United Kingdom
- au - Australia
- ca - Canada

Domain Names

A domain name can have several parts. Unique domain names mean that multiple sites can have individual computers with the same local name. When used, an Internet address is translated to an IP address by software called the Domain Name System (DNS). There is no one-to-one correspondence between the sections of an IP address and the sections of an Internet address.
The World Wide Web

- The World Wide Web allows many different types of information to be accessed using a common interface.
- A browser is a program which accesses and presents information.
- A Web document usually contains links to other Web documents, creating a hypertext environment.
- The term Web comes from the fact that information is not organized in a linear fashion.

Web documents are often defined using the HyperText Markup Language (HTML).

Information on the Web is found using a Uniform Resource Locator (URL):
http://www.bilkent.edu.tr/information/student.html
http://www.google.com
ftp://java.sun.com/applets/animation.zip

A URL indicates a protocol (http), a domain, and possibly specific documents.

Problem Solving

- The purpose of writing a program is to solve a problem.
- The general steps in problem solving are:
  - Understand the problem
  - Dissect the problem into manageable pieces
  - Design a solution
  - Consider alternatives to the solution and refine it
  - Implement the solution
  - Test the solution and fix any problems that exist

Many software projects fail because the developer didn't really understand the problem to be solved.

We must avoid assumptions and clarify ambiguities.

As problems and their solutions become larger, we must organize our development into manageable pieces.

This technique is fundamental to software development.

Programming Languages

- A programming language specifies the words and symbols that we can use to write a program.
- A programming language employs a set of rules that dictate how the words and symbols can be put together to form valid program statements.

Language Levels

- There are four programming language levels:
  - Machine language
  - Assembly language
  - High-level language
  - Fourth-generation language

Each type of CPU has its own specific machine language.

The other levels were created to make it easier for a human being to read and write programs.
Basic High Level Languages

- Imperative Languages
  - Fortran, Algol, Cobol, Pascal, C, Ada

- Functional/Logic Languages
  - Lisp, ML, Miranda, Schema
  - Prolog

- Object-oriented Languages
  - Smalltalk, C++, Java

Programming Languages

- A program must be translated into machine language before it can be executed on a particular type of CPU
- This can be accomplished in several ways
  - A compiler is a software tool which translates source code into a specific target language
  - Often, that target language is the machine language for a particular CPU type

Syntax and Semantics

- The syntax rules of a language define how we can put together symbols, reserved words, and identifiers to make a valid program
- The semantics of a program statement define what that statement means (its purpose or role in a program)
- A program that is syntactically correct is not necessarily logically (semantically) correct
- A program will always do what we tell it to do, not what we meant to tell it to do

Errors

- A program can have three types of errors
  - The compiler will find syntax errors and other basic problems (compile-time errors)
    - If compile-time errors exist, an executable version of the program is not created
  - A problem can occur during program execution, such as trying to divide by zero, which causes a program to terminate abnormally (run-time errors)
  - A program may run, but produce incorrect results, perhaps using an incorrect formula (logical errors)

Basic Program Development

- Edit and save program
- Compile program
- Execute program and evaluate results

Summary

- Chapter 1 has focused on
  - components of a computer
  - how those components interact
  - how computers store and manipulate information
  - computer networks
  - the Internet and the World Wide Web
  - programming and programming languages