Computer Basics Foundation

TEIA

Student Learner's Guide

Computer Basics Foundation Courseware
Written by Kelvin MacDonald
Published by Velsoft Training Materials Inc.

Courseware Release Version 3.0

Notice of Rights

Trademark Notice

Terms such as PowerPoint, Windows, Word, Microsoft, etc. are trademarks of Microsoft, Inc. Throughout this courseware title, trademark names are used. Rather than just put a trademark symbol in each occurrence of a trademarked name, we state we are using the names only in an editorial fashion and to the benefit of the trademark owner with no intention of infringement of the trademark.

Notice of Liability

The information in this courseware title is distributed on an 'as is' basis, without warranty. While every precaution has been taken in the preparation of this course, neither the authors nor Velsoft Training Materials, Inc. shall have any liability to any person or entity with respect to any loss or damage caused or alleged to be caused directly or indirectly by the instructions contained in this book or by the computer software and hardware products described in it.

Disclaimer

We make a sincere effort to ensure the accuracy of the material described herein; however, Velsoft Training Materials, Inc. makes no warranty, expressed or implied, with respect to the quality, correctness, reliability, accuracy, or freedom from error of this document or the products it describes. Data used in examples and sample data files are intended to be fictional. Any resemblance to real persons or companies is entirely coincidental.

All information in this manual was correct at the time of writing. We are not affiliated with nor have any control over changes made to the product described in this manual. These include, but are not limited to, changes in the application's color scheme, icon appearance and locations, addition or removal of program features, online templates, and help content. We reserve the right to make corrections to the courseware at any time and without notification.

Table of Contents

SECTION 1: GENERAL CONCEPTS

In this section you will learn:

- What a computer is
- What computer hardware is
- What software is
- What peripheral devices are
- What information technology means
- What a personal computer is
- What a main frame is
- What a network computer is
- What a laptop is
- What a PDA is
- About the CPU
- About RAM
- About the power supply and cooling system
- About disk drives
- About input/output devices
- About ports
- How CPU speed affects performance
- How RAM affects performance
- How applications affect performance
- How other factors can affect performance

Lesson 1.1: Basic Terms

When you are learning about computers, or any other subject for that matter, it is always a good idea to start with a solid foundation. In keeping with this idea, these early lessons will introduce you to some of the most basic concepts regarding computers and computer related technology.

In this first lesson, you will learn what a computer is. You will also receive a gentle introduction to the basic concepts of hardware, software, and peripheral devices. Once these topics have been covered, the lesson will be closed with a discussion of what is meant by Information Technology.

What is a Computer?

Nowadays, computers come in many forms. When you say the word "computer," the image that most readily springs to mind is a rectangular box, a screen, and a keyboard sitting on top of a desk. While the computer has indeed become a common fixture on desktops all over the world, they are also frequently found in the form of laptops and handheld devices.

A less obvious but extremely common application for computers can be found in their use as embedded devices in a multitude of electronic products, like cameras, mobile phones, toys, appliances, etc.

Even though computers come in many different shapes and sizes, there are some general ideas that they all have in common. A computer is a machine that processes (manipulates in some way) data or information according to a sequence of instructions. The ability to store and execute sequences of instructions is very important as it is this ability that allows computers to carry out different tasks.

A computer can be described (in a very fundamental way) as programmable machine that can store and execute a sequence of instructions. However, while this definition may be technically accurate, most people use the word computer in a broader sense that refers to the programmable machine, its storage devices, its monitor, keyboard, mouse, and a host of other attachments.

Beyond describing a computer with a definition of what it is, you can also describe a computer in terms of what it can help you do.

On a personal level, a computer can be used to store records and files, communicate locally or internationally, prepare a variety of documents and reports, edit photos and video, and provide you with access to a global network of information. Beyond this, computers are now used in science, health, business, and government to such an extent that it would be extremely difficult to function without them.

What is Hardware?

As you now know, a computer can be described as a programmable machine that can store and execute a series of instructions. A computer's hardware consists of the actual physical components that are used to store and run the programming instructions. The wires, circuits, and microchips that a computer is built from can all be called computer hardware.

The physical storage devices (machines used to store and retrieve data) and input output devices (like keyboards, monitors, and printers) are also computer hardware components.

Basically, any tangible part of a computer system (a part that you can see and touch) is considered hardware. Furthermore, pretty much any computer hardware component will fall under one or more of the following four categories.

Storage/Memory	The parts of a computer where programs and data are kept.
Input	Devices that help convey data into the computer, such as a scanner, keyboard, mouse, or microphone.
Output	Devices that help the computer output information to the users, such as a printer, monitor, or speakers.
Control/Processing	These very important hardware components interpret the sequence of instructions in the computer's running program. Electronic signals (based on these program instructions are then sent to the other hardware components to control their actions. A computer's Central Processing Unit (often called a CPU or processor) is a good example of a piece of control /processing hardware.

On a final note, it is important to understand that some pieces of computer hardware can simultaneously belong to more than one of the categories listed above. For example, a network interface card (a component that helps facilitate communication between different computers) could be thought of as both an input device and an output device.

What is Software?

A computer program is an organized set of instructions that is designed so that it can be stored and executed on a computer. Collectively, computer programs are often referred to as computer software.

A modern computer can run many different programs, with each program having its own specific purpose. As each instruction in a program is executed, the computer's control hardware sends messages to the other parts of the computer to cause them to function as directed by the instructions. Unlike hardware components, which are physically tangible ("hard") objects, computer (software) programs are stored and run as electrical signals inside the computer.

Even though you cannot reach out and touch them, these software programs are an essential part of any computer system. It is the software that tells the computer hardware what to do and how to behave in response to a user's actions. Without software, a computer is little more than a box of complex, but useless, electrical parts.

Software can be written to perform an amazing variety of different tasks on a computer. The fact that many different software programs can be loaded and run on a single computer is what makes modern computers so versatile. Today, a typical office computer can be used for word processing, accounting, sending e-mail messages, and more, depending on what software has been installed.

To summarize, software programs are sets of instructions that are executed on a computer. These instructions are processed in the form of electrical signals by the computer hardware to provide functionality to the computer system.

Here are some helpful insights and information about modern computer software.

- Generally software can be organized in terms of two broad categories: applications software and systems software.
- **Applications software** refers to programs that are designed for a relatively specific subject area or purpose, like word processing, e-mail, web browsing, spreadsheets, databases, animation and graphics, and so on.
- **Systems software** generally refers to software programs that are designed to provide and maintain an environment that allows the user (and the applications software) to interact with the computer hardware. Operating systems like Windows, Linux, and UNIX are a few examples of systems software.
- Software can be stored for long periods of time on a variety of different media, including DVD's, hard disks and USBs. The storage media is constantly changing as technology brings in greater storage capacity and new ways of recording data.
- Software can be temporarily stored in a computer's main memory (RAM).
- Software was often published to, and installed from, optical disks like DVD's but the more common method now in use is where the software can be obtained over network connections, using a method called downloading.

What are Peripheral Devices?

A peripheral device is a piece of hardware that is connected to a computer to expand a computer's functionality, e.g. if you want to create paper copies of the work you do on a computer, you could connect a printer. The printer is a peripheral device that, when connected to your computer, can produce printed (hard copy) pages of specific information stored on your computer.

Some other examples of peripheral devices are microphones, speakers, scanners, digital cameras, USB or flash drives, and keyboard/mouse devices. Just about anything that connects to a computer to improve or expand its functionality can be referred to as a peripheral device.

What is Meant by Information Technology?

Information technology can be described as the design, development, and implementation of information systems especially about the use of computer and telecommunication technology.

Essentially, Information Technology (also referred to as IT) deals with all aspects of information storage, retrieval, transformation, communication, security, and accessibility. Of course, because computers are such great tools for handling information, computer hardware, software, and networking are key components of the Information Technology field.

When people speak about occupations in the IT (Information Technology) field, they are often referring to jobs that involve computer networking, network administration, software development, technical support, Internet services, and web development.

Lesson 1.2: Types of Computers

In Lesson 1.1, we discussed some of the most basic computer related concept, including the question, “What is a computer?”

In this lesson, we will follow up on that basic question by discussing some of the main types of computers that are in use today. These computer types include: the personal computer, the mainframe computer, networked computers, laptop computers, and PDAs (Personal Digital Assistants).

What is a Personal Computer?

A personal computer is a small, affordable computer that is intended for personal or individual use. Typically, the main element of a personal computer system consists of a box that contains the main memory, hard drives, processor (CPU), circuit boards, and related devices required to run software programs.

In addition to this box, a personal computer system will usually require a keyboard, a monitor (screen), and a mouse (pointing device). Often, owners of personal computers will acquire additional peripheral devices, such as printers or scanners.

Personal computers are often called desktop computers because their size, computing power, and affordability have made them a common fixture on home and office desktops around the world. The familiar view of a rectangular box, with a screen and keyboard sitting on a desk, is what many people envision when they think of a computer.

Mainframe Computers

Mainframe computers are found at the opposite end of the computer spectrum from the PC (personal computer). Unlike a personal computer, a mainframe computer is large and expensive. A mainframe may be larger than a refrigerator in size and can cost several millions of dollars.

In general, a personal computer is designed around a single processor (CPU) and is used primarily on an individual basis in an office or home setting.

A mainframe, on the other hand, can be designed around multiple CPUs, and can support multiple (sometimes hundreds) of users. Mainframes also have vast storage capabilities, far exceeding that of a typical personal computer. Mainframes are typically purchased by large companies and institutions to help manage immense amounts of important data. Because of this, mainframe computers are designed to be very fault tolerant and can often be serviced without being shut down. As a matter of fact, it is quite common for mainframe computers to be up and running for years at a time.

Another interesting feature of mainframes is that most can run (host) several operating systems simultaneously, whereas a personal computer will typically only be running a single operating system at any given time.

What is a Laptop?

A laptop is a small (usually between 1 and 5 kilograms in weight) portable computer that can be operated on battery power for a period without any dependence on an external power supply.

This makes laptop computers a great choice for users who want the processing power of a personal computer, but also require the convenience of a portable, lightweight machine that does not always need to be plugged in. In this regard, laptops are quite versatile, and can be used in cars, on trains, in classrooms and offices.

The hardware parts inside a laptop are quite similar in purpose to the parts of a regular personal computer; however, laptop components are usually smaller and are designed to require less power.

When a laptop's battery loses power, the laptop can be plugged into an external power source by using an external AC/DC adapter. When plugged in, the laptop is provided with enough power to run normally and charge its battery.

Typically a laptop computer will have a slim screen that folds neatly down to reduce its overall profile. In addition, a laptop has a built-in keyboard and a touchpad or another pointing device. Most laptops also support the attachment of an external mouse (pointing device) as well.

As mentioned previously, a laptop can have comparable functionality (in terms of computing power) to a larger (desktop) personal computer; however, a laptop will usually be more expensive than its comparable desktop counterpart. Also, because a desktop personal computer has more room to accommodate hard disks, a typical personal computer will often have a greater secondary storage capacity than a laptop.

Laptops are also sometimes referred to as notebook computers.

Lesson 1.3: Anatomy of a PC

In this lesson, you will start to get an idea of what is inside the box of a typical desktop PC. You will begin this lesson by learning about the CPU, which is the "brain" of a personal computer and a major factor in a computer's performance capabilities. Following this, you will be introduced to the concepts of main memory (RAM), buses, input/output devices, and other internal components like fans and the power supply.

The CPU

In a personal computer, the CPU (Central Processing Unit) is responsible for interpreting program instructions, performing calculations, and directing other components. In a sense, the CPU can be thought as the brain or control center of the computer.

A central processing unit can itself contain internal building blocks that handle different types of jobs or tasks. For example, a typical CPU will have an ALU (Arithmetic Logic Unit), a CU (Control Unit), CPU registers, and usually a high-speed local storage area called a cache.

The CPU's control unit interprets the instructions of a program and signals/controls the other computer components based on the instruction.

The CPU's arithmetic logic unit is used to perform calculations (arithmetic) and logical operations (i.e. AND, OR, EXCLUSIVE OR, NAND) on values. The arithmetic logic unit (ALU) and the control unit (CU) of a CPU are extremely important factors influencing how the CPU will perform.

The CPU registers are very fast storage areas that often hold the program instruction that is currently being processed (interpreted by the control unit). Typically, a program instruction is loaded from memory into a CPU register at which point the given instruction will be executed by the CPU.

A CPU cache is a fast memory storage area that holds values that are frequently requested from the computer's main (but slower) memory. Because CPU cache memory is faster than main memory, a speed advantage is gained by storing frequently accessed memory values there. The CPU can quickly access the data in its cache, rather than retrieve the same data from the slower main memory every time it is needed.

In a personal computer, the ALU, CU, onboard cache, and registers are usually all designed into a single chip. This chip is attached ("plugged in" so to speak) to a motherboard (circuit board) to facilitate communication with other computer components like input/output devices, main memory, and secondary storage (hard disks).

These single chip CPUs are also called processors and come in a wide variety of models with different features and speeds. A couple of well-known manufacturers of CPUs are Intel and AMD (Advanced Micro Devices).

RAM

To execute software programs, a computer must be able to calculate values, perform logical operations, and store (remember) instructions or data. As you have seen, the CPU is used for calculating values and performing logical operations. To a small degree, a CPU can even handle some storage tasks by using its registers and memory cache.

The problem is that modern computers are often required to run huge software programs and deal with vast quantities of data. To meet these demands, additional memory storage is required.

Most personal computers contain a large amount of RAM to meet the computer's working memory needs. RAM stands for Random Access Memory, which means that each individual memory location (the place where an instruction or data item is stored) can be accessed in the same amount of time. All memory locations can be accessed in the same amount of time.

Typically, RAM is attached (plugged in) to the PC's motherboard in the form of modules. Different modules can have different access speeds and different memory capacities. The instructions for active (running) programs and the data required by these running programs are often stored in the RAM. In this sense, you can say that RAM provides the working memory that a computer needs to store the data and instructions for the programs that are currently running.

There are a few more important things to remember about your PC's RAM. RAM is volatile, meaning that whatever is stored in RAM will be lost when the computer is shut down. This is because RAM needs a small amount of electrical power to preserve the data and instructions that it stores.

RAM is a lot faster than secondary storage devices like hard disks, USBs and optical drives. This is part of the reason that modern computers can perform so well; the more RAM that a computer has, the faster storage is available for programs and data.

RAM is slower than cache memory or CPU registers, but it is also less complex and less expensive. Therefore RAM is used as the main working memory of a PC, while cache memory is used for quick access to a relatively small quantity of instructions or data.

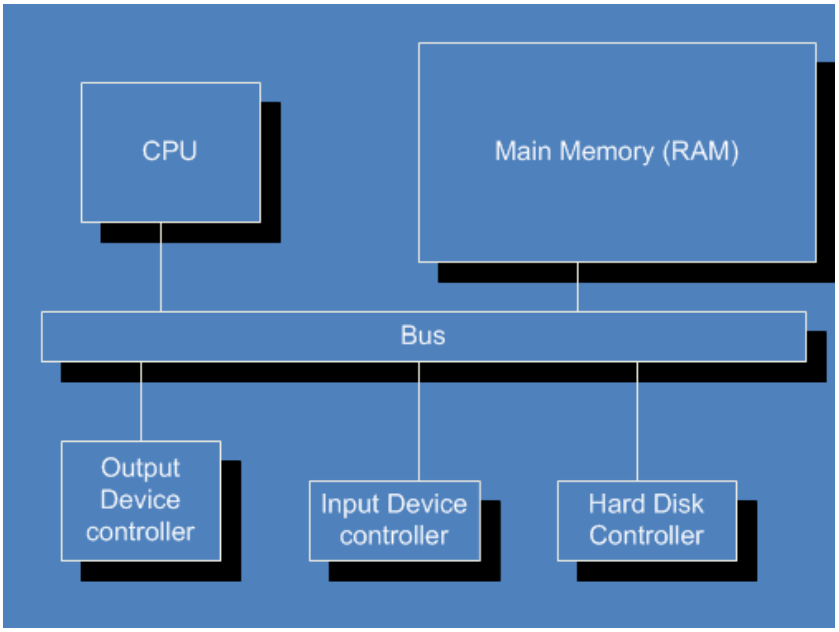
Buses

Simply put, a bus is a transport route for sending information or power between computer components. It is sometimes useful to think of a bus as a parallel series of wires, running between your computer components. These wires can connect the different components so that they can communicate and exchange instructions/signals, and data to each other. (The actual physical construction of a bus does not have to resemble a series of parallel wires.)

Imagine a very simple PC design as an example. In our imaginary PC, we have a CPU, some modules of RAM memory, and some input/output devices. For a program to run, instructions must be fetched from the computer's RAM and sent to the CPU.

An instruction that is processed by the CPU may request data or another instruction, in which case that data or instruction will also need to be retrieved from the RAM. The instruction may also direct the CPU to send data to the RAM memory for storage, or to send data to the input/output devices for the user.

All these components must be able to communicate and transport information in some way for even this simplistic computer to function.



As you can see from the diagram of our imaginary PC, the CPU, the RAM, and the input and output devices can send and retrieve information in the form of electrical signals passed over buses.

The real internal architecture of even a simple PC is somewhat more complex than the description and diagram given above. The important point is to understand that components like the CPU, RAM, and various input/output devices rely on buses to communicate and transport data and/or instructions.

Input and Output Devices

Computers would be of very little use to us if we could not communicate with them. To operate a computer, a user must be able to input data and commands, and moreover, see or hear the results that the computer produces.

The devices that are used to feed information into a computer are called input devices. The devices used by a computer to present information to the user are called output devices.

Of course, computers can get information from other sources like hard disks and USBs, but these are generally considered to be storage devices rather than input devices.

The main input devices employed by most PC users are the mouse and keyboard.

The keyboard allows the user to enter uppercase and lowercase letters, numbers, familiar symbols, and commands. With a keyboard, you can type sentences, mathematical equations, and enter data into a variety of computer forms and applications. Keyboards are used for an extensive list of tasks, including but not limited to word processing, computer programming, composing e-mail, entering data into spreadsheets, and issuing commands.

A mouse is a type of pointing device that translates movements and actions made by the user into the movements and actions of a small symbol (mouse pointer) on the computer screen. Computer mice have buttons (normally at least two) that can perform various actions when clicked. This means that a mouse can be used to point to objects on the screen and even perform actions on these objects. Frequently computer mice are used for dragging objects (moving them), opening objects (displaying them), or starting or closing software programs. The limitations of what can be done with a computer mouse are mostly imposed by the type of graphical user interface that is available. In most situations, anything that does not involve the typing of numbers, text, or symbols, can be done with a mouse.

Another type of input device is a scanner. These devices can be used to convert an ordinary photograph, drawing, or sheet of text into a digital image. The object that you want to scan is placed on the scanner, the scanner is activated, and the object (picture, drawing, or a handwritten or typed page) is converted into a digital image that can be stored on your computer.

A computer printer is an output device used to produce a permanent, printed (paper) copy of an image, text document, or other output file from your computer. The user indicates what file or document they want a printed copy of (usually by using a mouse), and then uses a print command to send the specified file or data as output to the printer. At this point, the printer produces a copy of the specified information on paper.

A computer screen, or monitor, is the output device that users rely on most. There are many different sizes of screens available for a personal computer, but larger, higher quality screens are

generally more expensive than smaller ones. Nearly all modern computer monitors have full color displays and support a selection of different screen resolutions.

Human computer interaction is made possible through input and output devices. This is how people and computers communicate. Devices like a mouse, a keyboard, and a monitor are used so frequently that they are now considered by most people to be essential components of a personal computer system.

Disk Drives

Basically, computer disk drives come in two types: flash drives and magnetic drives. Magnetic disk drives include hard disks and floppy drives.

The main difference between the two drive types lies in how each type stores its data. With flash drives the memory is a type of electronically erasable programmable read-only memory (EEPROM), memory chips that retain information without requiring power. (This is different from flash RAM, which does need power to retain data.) Inside the flash chip, data is stored in cells protected by floating gates. Tunneling electrons change the gate's electronic charge in "a flash" (hence the name), clearing the cell of its contents so it can be rewritten.

With magnetic disks, signals are stored as tiny, magnetized areas on a rotating platter (for a hard drive) or on a thin magnetized film. When these magnetic surfaces pass under a read/write head, the magnetic signals can be detected and sent as data to the computer.

The main difference between hard disk drives (HDDs) and flash drives is the lack of moving parts, and the reduced risk of physical damage in the flash drives because of this. There is also a difference in the way flash drives writes and stores data; hard disk drives write data to platters in a linear fashion, to sectors of a standard size. USB flash drives, however, spread the individual pieces of data that form files all over the chip, and pieces of data can be found in several locations within a flash drive.

A modern hard disk is really a stack of disks (platters) that form a kind of cylinder. Each disk in the cylinder has a magnetic coating which allows data signals to be stored on it. Since data can be stored on both sides of a platter, and there are multiple platters in a cylinder, a typical hard disk can hold quite a bit of information.

In a PC, a hard disk is used to store the computer’s operating system, application programs, and any other files or data that the user wants to retrieve quickly and easily.

Power Supply and Fans

Nearly every component t inside a PC, from a hard disk to a CPU, requires electrical power to work. These components require DC (Direct Current) electricity while the power that is available at a typical wall outlet is AC (Alternating Current).

To confound matters even more, different PC components require different amounts of electrical power. For example, a CPU may require 30 or 40 watts of power, while a hard disk drive may require only 5 or 10 watts. If a PC has a lot of components, like multiple drives, fans, a CPU, and

various graphics or network cards, the power required by these devices adds up. Therefore computers contain an important device called a power supply.

A power supply is a box like component that rests inside a PC. Looking at the back of a PC, you can often tell where the power supply is located by the presence of a power cord/wall plug, a power switch, and a fan.

Power supplies can generate a lot of heat as a byproduct of converting the AC current from the wall outlet to the DC current that is used by the computer hardware. If a power supply overheats, it could fail, rendering the PC useless. A fan integrated with the power supply helps to circulate air and dissipate heat, to help ensure that the power supply remains adequately cooled.

Typically, a PC power supply will have several small wires that end in connectors. When a PC is put together, the appropriate connector from the power supply is used to connect to the corresponding hardware device inside the PC. The types of power supply used in a PC will depend on the type of motherboard being used. (A motherboard is the main circuit board, containing buses, which the PC's internal components plug into to exchange data and information.)

If a PC has a lot of hardware components (multiple hard disks, drives, network interface cards, etc.), it is important for the power supply to be able to produce enough wattage to support all these devices.

Remember, any device that is inside the box in a PC (any device that does not plug into the wall on its own) will get its power from the PC's power supply. Peripheral devices (like printers and scanners) and output devices (like speakers or monitors) have their own power cords that plug into a wall outlet.

Fans are important to PCs because of the cooling effect they provide. If a power supply overheats, it can fail. Similarly, if the PC's CPU overheats, it can cause errors in program execution, or fail entirely.

To prevent overheating, CPUs often use a physical cooling device (analogous to a car radiator) to help conduct heat away from the CPU itself. These devices are frequently combined with small CPU fans that help to further dissipate heat. Most PCs will have a fan incorporated with the power supply, as well as a CPU fan to help cool the CPU directly. In some PCs, fans are also used to help with the general circulation of air through the PC's enclosed box. When air circulates freely through the PC, heat is more easily dissipated from the internal components.

More on Peripheral Devices

Peripheral devices are additional (non-essential) components that can be added to a computer to provide more functionality.

Devices like external drives, printers, speakers, scanners, lash drives and digital cameras that can connect to a PC, but that are not essential to its core functionality, are peripheral devices.

As you already know, computers can output data to a printer for a permanent hard copy, and they can output video data to a screen so the user can see and interact with files and software

programs. Not surprisingly, computers can also generate, process and output audio data. Note, most computers come with a basic internal speaker, but this is generally of low audio standard and most users seeking audio will add external speakers.

Computer speakers are typically connected to a computer by plugging into the computer's sound controller (sound card) with either a stereo jack plug or a USB plug into a PC. Often computer speakers will come with their own controls for adjusting the audio tone (base, treble) and volume. Users can also control the volume and tone of the audio output by using software programs on the PC.

Speakers are important if the user wants to be able to hear the audio part of multimedia files. If a user wants to play media types such as audio files or various video file formats, they should have speakers.

Often various Web sites will feature streaming audio and video that can be accessed with an Internet connection. Once again, speakers are required for the user to hear the audio output.

Hardware Ports

Any peripheral device or input output device must be connected to a computer in some way. To support the wide variety of devices that are available, and to provide backward compatibility with older devices, a variety of hardware ports are usually available on the average PC.

On the back of a PC are different types of ports that devices may be plugged into. Commonly used hardware ports like USB are often located on the front of the PC for easy access.

The following table provides a list of some of the most common hardware ports on a PC, with a brief description of each.

LPT (Parallel Port)	LPT ports look fairly large (long) compared to other ports, and they contain 25 small holes, designed to receive the pins at the plug end of a printer or scanner's connecting cable. For the most part, these ports are used to connect printers and scanners, and they are commonly referred to as printer ports. Data transfer rate for traditional LPT port is up to about 150 kilobytes per second. However, the newer enhanced parallel ports (EPP) support higher data rates of 500 kilobytes/sec, up to 2 megabytes (2000 kilobytes)/second.
Serial Ports	A serial port typically has 9 (but sometimes 25) small metal pins. A device that connects to this port will have a plug with holes designed to receive the pins. Serial ports differ from parallel ports in that they exchange information one bit at a time, rather than several bits at a time in parallel.

Serial ports were most often used to connect to external modems, and sometimes for connecting to a mouse. In recent times, serial ports are not used as much. In very new PCs, they may not even be present.

Data transfer rate for serial ports is typically less than 60 kilobytes per second.

USB Ports

USB stands for Universal Serial Port, a type of port that is commonly used in current generation PCs. These ports have a small rectangular socket that the connecting wire from the USB device plugs into.

USB ports can be used to connect to a wide variety of devices like cameras, external flash drives (and other drive types), mice, keyboards, printers, and more.

Because USB ports are so widely used, they are rendering serial ports (and to an extent, LPT ports) obsolete.

Currently, there is a third generation USB port (USB 3.0) available on PCs which offers higher data transfer rates than the first and second USB versions. Fortunately, USB devices that are designed for the original USB ports will work with the new USB ports (but at a slower data rate). Also, devices designed for the new USB ports will work with the older style USB ports (but also at a lower data transfer rate).

PS/2 Ports

A PS/2 port is used to connect a mouse or keyboard to your computer. These ports have small circular socket that contains smaller holes inside to receive pins from the plug end of the connecting device. These are generally now obsolete having been replaced with the USB ports.

DB9-15 Video Port

This port is used to connect a computer screen (monitor) to your PC. The port has 15 holes (three rows of five each) that are designed to receive the pins from the monitor's connecting wire.

A 15-pin video port looks a lot like a serial port, except that the video port has holes whereas the serial port has pins.

1394 FireWire Port

These ports look like small rectangular sockets, but they are squarer looking than USB ports.

FireWire ports can be used to transfer large amounts of data at a high rate of speed. They are often used to connect to such peripheral devices as digital video cameras, and external hard drives. Again these ports are not as common as they once were

	due to the increasing speeds of USB.
Ethernet (network) Port	An Ethernet port is a rectangular socket that looks a lot like the type that a phone jack is plugged into, but a little bigger. Network cables can be plugged into this port to provide access to high-speed Internet modems or local area networks. Ports of this type provide can provide high data transfer rates across networks in the high megabits and low gigabit per second range. The transfer rate that an Ethernet port provides depends on the type of network interface card being used and the overall speed of the network itself.
HDMI port	A High Definition Multimedia Interface (HDMI) port is designed as a digital interface used to receive and transmit both video signals and audio signals from an array of source devices through to several display monitors. In many instances this has replaced the Firewire ports.

Lesson 1.4: How a PC Works

In this lesson, you will learn about CPU speed, how RAM impacts on performance, how software applications affect performance, and how other factors like your hard disk or network connection speed can affect performance.

CPU Speed and Performance

The two main factors that have the most impact on computer performance are speed and storage. As such, you will begin this lesson by learning how CPU speed affects performance.

The CPU (processor) serves as the central manager or “brain” for a PC. It should come as no surprise that the faster a CPU works the better a PC will perform.

CPU speed is often specified in terms of something called clock speed. Each CPU has an internal clock that ticks at a certain rate. Any given CPU will require a certain number of ticks to execute an instruction. Since computers run a program by executing the sequence of instructions that is specified in the program’s code, the faster an instruction can be executed, the better the computer will perform.

Clock speeds are measured in terms of Hertz (Hz). 1 Hz means one complete cycle per second. When something is so fast that it can perform one million cycles in a second, you can say that it has a speed of 1 Megahertz (MHz). If something was so fast that it could complete one billion cycles in a second (1000 megahertz), you could say that it has a speed of 1 Gigahertz (GHz).

If a CPU has a clock speed of 500 MHz, it can perform 500,000,000 clock cycles (ticks) in a second. If this same CPU takes 100 ticks (clock cycles) to execute an instruction, it could (in theory) perform 5,000,000 instructions per second.

Now suppose we have another CPU that also needs 100 clock cycles to execute an instruction. If this CPU has a clock speed of 1 GHz, it will be able to perform 1,000,000,000 clock cycles in a second. At 100 cycles per instruction, it would be able to execute 10, 000,000 instructions in a second.

In general, the CPU clock speed is a good indication of how fast a given processor can execute instructions. However, CPU clock speed is best used to compare CPUs from the same basic family. This is because different CPU architectures (designs) can require different amounts of clock cycles to execute a single instruction. It can be possible for a CPU that can execute an instruction on only a few (or one) clock cycle to outperform a CPU with a higher clock speed, but that also requires many clock cycles to execute a single instruction.

In today's PC, CPUs with speeds of above 3+ GHz (gigahertz) are not uncommon.

RAM and Computer Performance

The amount and type of RAM (main memory) used in a computer can also affect the computer's performance. RAM is volatile memory that the computer uses to temporarily store the programs and data that are in use currently. Because the processor may frequently request that data be moved to and from the RAM, the speed at which the RAM can respond to these requests is important.

Data in RAM can be accessed (read or written) much faster than the data on a hard disk. Random access memory is made of integrated circuits (chips) and therefore does not require moving mechanisms like read and write heads that scan over rotating disks.

The actual speed of RAM can depend on the type of RAM used and the speed of the connecting buses that the data must travel when the RAM is accessed. The two main categories of RAM are static RAM (SRAM) and dynamic RAM (DRAM).

DRAM is the most common and inexpensive type of RAM found in a typical PC. This RAM is made up of a vast number of tiny capacitors. These tiny capacitors can store electrical information which can be interpreted by the computer as values. Because a capacitor loses the electrical power that it holds over time, the many capacitors in a DRAM memory module must be refreshed with electricity at regular intervals to prevent information loss. For this reason, this type of RAM is said to be dynamic, or DRAM.

Think of SRAM as being made from many tiny switches that can store data. Because the switches do not need to be refreshed with electrical power once they have been set, the information in SRAM is static (in other words, it does not have to be refreshed like DRAM).

SRAM is faster than DRAM, but it is also more complex and expensive. For this reason, it is used primarily for external (outside the CPU) cache memory.

DRAM can be made very dense (lots of memory storage in a small size) and it is also relatively inexpensive. For these reasons, it is used for the vast majority of RAM storage in a typical PC.

The access speeds of RAM are measured in nanoseconds (billionths of a second) whereas the access speeds of a hard disk drive are measured in milliseconds (thousandths of a second). This means that RAM is on the order of a million times faster than a hard disk.

Another aspect of RAM that affects your computer's performance is the quantity that you have available. Basically, programs and instructions that are stored in RAM do not have to be loaded from the much slower hard disk drive. The more information you can fit into your RAM, the fewer hard disk accesses (which are comparatively very slow) will be required.

With fewer hard disk accesses, your overall computer performance will be much better. This is because the computer will be accessing storage in nanoseconds (RAM speeds), instead of milliseconds (hard disk speeds).

How Applications Affect Computer Performance

The type of applications that you run can also affect a computer's performance. Different types of applications make different demands on a CPU and RAM. Some software applications, like a word processor, are I/O bound (input/output bound). This means that the rate the computer works on a task is dependent on the user.

With a simple word processing program, the computer spends most of its time waiting for the user to type characters and view what is being output to the screen. Though these waits do not seem like much to the user, a one GHz CPU might be able to perform millions of instructions during these idle times. So, while waiting for the user, the CPU has a lot of clock cycles that it can use to work on other software programs that may also be running.

Some programs, however, can be CPU bound. This means that these programs involve a lot of heavy calculation (number crunching) more than user input. When this type of program is running, the CPU may not have as much idle time to spend on other running programs.

As an example, a user may be able to use a word processor, a spreadsheet program, surf the web, and listen to an MP3 file all at the same time without noticing any decline in performance.

If, on the other hand, a person was using a program that converts between digital video formats while listening to an MP3 audio file and playing a graphically intense video game, they may notice a decline in the computer's performance.

In addition to the type of software that is running, the amount of software that is running is also important.

Each software program that is running requires some space in main memory (RAM). If there are a lot of different software applications that require space in RAM, there will be less space available for each. This means that there will be more frequent hard disk accesses as information is exchanged between main memory and the hard disk drive. More hard disk accesses lead to more waiting by the applications for the instructions or data that they need.

If software applications are frequently accessing a hard drive, the performance of the computer may be sluggish.

Other Factors

There are some other factors that can affect the performance of a computer system, e.g. the speed at which the buses can transfer information can be an important factor. The bus's speed will depend on its width and its design architecture.

A larger amount of cache memory (high speed SRAM) and CPU cache (onboard the processor) can also improve a computer's performance as more data or instructions will be stored in these fast access areas.

Hard disk drives with very high rotation speeds and reduced seek times can help reduce hard disk access times and thereby improve overall system performance.

There are also other components like video (graphics) cards that can help computer performance for applications that involve a lot of video/ graphics processing. These cards often contain their own RAM and other hardware optimizations that reduce the workload placed on the CPU by intense graphics-oriented software (like some games).

Finally, not all software is created equal. There are often many ways to write a program that performs the same task, and some ways of writing software to perform a task are more efficient than others. Some software may even be optimized to take advantage of certain types of hardware (as is often the case in video game consoles).

SECTION 2: HARDWARE DEVICES

In this section you will learn about:

- What the CPU does
- How CPU speed is measured
- What RAM and ROM are?
- How computer memory is measured
- The keyboard
- The mouse
- The scanner
- The microphone
- Other input devices
- The computer screen
- Touch screens
- Printers
- Plotters
- Speakers
- Internal and external hard drives
- Floppy disks
- CDs and DVDs
- Zip disks
- USB drives
- Data cartridges and tape drives

Lesson 2.1: CPU and Memory

In this lesson, you will learn more about the CPU and what it does. You will also learn how CPU performance is measured, and about the differences between RAM and ROM. Finally, to close the lesson, you will learn how to understand the units used to measure storage capacity and file size.

What is a CPU?

The CPU is the control center and main workhorse for a computer. This single component is responsible for executing instructions from software programs, and then directing the other parts of the computer based on these instructions. The CPU also handles logical operations and mathematical computations, serving as the “brain” of the computer.

A typical CPU can be subdivided into different units that handle different tasks. For example, a typical CPU may have a control unit, an arithmetic logic unit, a data-path logic unit, CPU registers, and an onboard memory cache.

A powerful CPU can perform a lot of tasks at a very high rate of speed. All things being equal, the more powerful the CPU, the better the performance of the computer. Some computers (like mainframe computers) can have several CPUs running in parallel, to further increase performance.

What Does the CPU Do?

A computer is a programmable machine. This means that it can perform many different tasks depending on the instructions that are given to it. The main purpose of the CPU is to interpret these instructions and then control what happens next based on what the instructions specify.

To get a better idea of what a CPU does, let us take high level walkthrough of the execution of a simple software program.

A software program is a sequence of instructions for performing a specific task. When a software program starts, it is loaded into main memory (RAM).

To start execution, the first instruction from the program is retrieved from main memory and loaded into a CPU register in the control unit. This “instruction register” stores the instruction temporarily while the control unit processes it.

This instruction contains information that the CPU’s control unit can break down and interpret. Once the control unit interprets the instruction, it sends signals out that direct the other parts of the CPU (or other hardware components like input output devices) on what to do.

In a sense, you can think of the control unit as a kind of puppet master; pulling the strings that cause other components to act, based on the instruction that has been interpreted.

Once the control unit processes an instruction, a special CPU register called the program counter is incremented (its value is increased) so that it now holds the memory address (the address corresponding to a location in RAM) of the next instruction in the program.

The instruction corresponding to the memory location specified in the program counter is loaded into the instruction register, and the control unit interprets this next instruction to repeat the cycle. In this way, the sequence of instructions in the program is executed until the program ends or is halted by the user.

Sometimes, an instruction may specify that another instruction, other than the next one in sequence, be executed. In this case, the address of the new instruction is loaded into the program counter, and then this new instruction is retrieved from the memory location corresponding to that address. This means that program execution can skip to other instructions out of sequence if that is what the program instructions specify. This ability to jump around within the sequence of program instructions permits the use of subroutines and branches in the program.

Say that the control unit interprets an instruction that specifies a mathematical operation; for instance, two values are to be added. The control unit gets the memory addresses for the values from the instruction, and then has these values loaded from RAM into CPU registers.

At this point, the ALU (arithmetic logic unit) can take over as directed by the control unit and add the two values together. The result of the addition is placed in another CPU register, and then sent to the main (RAM) memory for storage. If needed, the ALU can perform other mathematical operations besides addition. It can also compare values and perform logical operations on them as well.

If a certain sequence of program instructions is repeated over and over (like a subroutine, for example) these instructions may be loaded into the CPU's cache memory. This will allow the control unit to access the instructions faster than if they had to be fetched from the main memory (RAM) every time.

Every software program that you run on your computer, no matter how complicated, must be executed as a sequence of instructions through a process something like the one just described. It is not hard to see how the speed of your CPU can have a clear effect on your computer's performance.

The following table briefly summarizes the CPU's components in the context of the previous discussion.

Component	**Role**
Control Unit	Interprets instructions and directs other components accordingly.
ALU	Performs arithmetic operations, comparisons, and logic when required by the program instructions.
CPU Registers	These are temporary storage areas for data (like memory addresses or mathematical operands) or for program instructions. CPU registers provide components like the control unit or ALU immediate access to the data and/or instructions that are currently being processed.

CPU Cache	The CPU Cache can be thought of as very high-speed RAM. It is used for quick access to frequently used instructions or data rather than fetching these instructions from the main memory (RAM).

CPU Performance Measures

Sometimes, CPU performance can be hard to gauge. Just because one CPU has a higher clock speed than another does not necessarily mean that it will perform better. CPU performance can be influenced by a range of factors, including the number of clock cycles required to execute an instruction, how complex a single instruction is, and the overall instruction set architecture of the CPU.

For example, a CPU with a CISC architecture (**C**omplex **I**nstruction **S**et **C**omputing) can perform a range of tasks (like loading a value into a register, and then loading another value into another register, and then adding the values together), by executing a single instruction. However, each complex instruction in a CISC CPU may take several clock cycles to complete.

On the other hand, a CPU with RISC architecture (**R**educed **I**nstruction **S**et **C**omputing) will perform only single or simple tasks with each instruction. For example, a RISC processor may require one instruction to load one value into a register, another instruction to load another value, and then a third instruction to add the values and place the result in a register. This means that a single instruction may not do as much (when compared to a CISC instruction), but it will probably take fewer clock cycles to perform.

Because of the possible differences in CPU architecture, using clock speed alone to gauge performance differences between CPUs from different families or manufacturers is not a great idea.

Luckily, there are other indications of CPU performance besides CPU clock speed. One such measure is the rate at which a CPU can execute instructions. The rate at which a CPU can execute instructions is measured in MIPS (millions of instructions per second).

If a CPU can perform 100 MIPS, the CPU can execute 100 million instructions per second. If one CPU can perform 150 MIPS, and a second CPU can perform 200 MIPS, the second CPU can perform 50 million more instructions in a second than the first.

At first, this seems like a good way to measure CPU performance, but upon closer examination, there are some drawbacks. As mentioned before, one CPU may be able to do more on a single instruction than another based on the CPU's instruction set architecture. This means that even though one CPU may perform more MIPS than another, there may not be that much difference in the actual work that gets done.

Another measure of CPU performance is called megaflops. When speaking about CPUs, a FLOP stands for floating point operation. Floating point numbers are numbers with fractional parts. For example, the number 3 is an integer (a whole number) and the number 3.1415926535 is a floating point number.

Operations involving floating point numbers require more work by the CPU than simple integer operations, and so as a measure of CPU speed, they can indicate how well the CPU will perform with heavy number crunching or calculating tasks.

The rate at which a CPU can perform floating point operations is typically measured in megaflops (Million **Fl**oating point **O**perations **p**er **S**econd). If a CPU can perform 100 megaflops, it can perform 100 million floating point operations in one second.

Though these measures may not be entirely good indicators taken on their own, they can be significant if taken together. This means that if one CPU has a higher clock speed, can complete more MIPS, and can perform more megaflops than another; it will probably have better overall performance.

It is important to remember that the CPU is just one component in a computer system, and overall computer performance can depend on many factors including the bus speed, how much RAM is available, how much cache memory is available, and even the speed of the hard disk drive.

RAM and ROM

RAM is volatile high-speed memory that the computer uses to temporarily store data and instructions. RAM is made of integrated circuits that contain millions of tiny storage capacitors. Electrical signals that can be interpreted as ones and zeros are stored in these capacitors as data. A single 1 or 0 is referred to as a bit, and a sequence of 8 bits is referred to as a byte.

RAM memory is broken up into addressable memory locations, with each location able to hold a single value or instruction. The CPU and other components can access the values (instructions or data) in RAM by referring to the memory address that corresponds to the stored value.

Some computers may have RAM that has an address resolution of 1 byte, meaning that each addressable memory location can store a single byte of data. Other computers can have an address resolution of a single “word” which refers to a specific number of bits. In many modern computers, the word size is 32 bits (four bytes), which means that RAM with an address resolution of a 32-bit word can store four bytes of data in each addressable memory location. Computers are also commonly available in 64-bit capacity.

An important feature of RAM memory is that it is rewritable, meaning that data can be stored in a memory location, and then overwritten in that same location with a new value. Also, any memory location in RAM can be accessed in the same amount of time as any other location in RAM.

There is also another type of memory called ROM (Read Only Memory) which is frequently used in computers.

ROM memory has the advantage of being nonvolatile; that is to say, it will not be erased when the power is shut off. On the other hand, it also has a disadvantage in that it cannot be easily rewritten (hence the term “Read Only Memory”). A ROM chip will always provide the same data or instructions every time it is accessed unless it is explicitly reprogrammed or rewritten through an intensive process.

Although some modern varieties of ROM can be rewritten (like flash ROM), it is generally a more time consuming and intensive process to write instructions or data to ROM than to RAM.

Because ROM is not easily or quickly rewritable, it is mostly used for storing data or instructions that very rarely need to be changed. RAM, on the other hand, is used for the opposite purpose; that is to say, RAM is used to store data and instructions that may have to change frequently and rapidly.

Measuring Memory

As mentioned above, information that is stored in computer memory can be thought of in terms of ones and zeros. A single one or zero is referred to as a bit, and a sequence of 8 bits is referred to as a byte. A single character like a letter or symbol will typically require a single byte of storage space.

A typical computer file, like a word processing document or a digital image, will require thousands of bytes of storage space. When we speak of thousands of bytes, we use the term Kilobytes (Kb). One kilobyte of memory will hold 1024 individual bytes. You will often see individual files on a computer, like documents or image files, that range in size from a few to several hundred kilobytes.

If a file is quite large (more than 1000 kilobytes) the size can be referred to in terms of megabytes. One megabyte is equivalent to 1024 kilobytes, or 1024 x 1024 =1048576 bits.

A typical computer folder that contains several files will often have a storage size of a few to several hundred megabytes.

If a file or folder is very large, taking up more than a thousand megabytes of space, we can refer to its size in terms of gigabytes. One gigabyte is equivalent to 1024 megabytes, or 1024 x 1024 kilobytes. Current generation computers often have up to gigabytes of RAM memory.

The following table summarizes these memory and storage space measurements.

Measurement	**Explanation**
Bits	Can store a 1 or a 0
Byte	Can store 8 bits. This is enough room to store a single letter or symbol.
Kilobyte (Kb)	Roughly a thousand bytes. Files like spreadsheets, word processing documents, and images are often a few to a few hundred kilobytes in size.
Megabyte (Mb)	Roughly a thousand kilobytes (or a million bytes). Folders and very large files may be several megabytes in size.
Gigabytes (Gb)	A gigabyte is roughly a thousand megabytes (or roughly a billion bytes). A small hard disk drive can store hundreds of gigabytes of information.

Terabytes (Tb)	A terabyte is roughly a thousand gigabytes (or roughly a trillion bytes). Some large capacity hard disks have terabytes of storage space.

Lesson 2.2: Input Devices

In this lesson, we will take a closer look at some of the more common input devices, e.g. the keyboard, the mouse, the scanner and the microphone.

The Keyboard

One of the most common devices that people use to input data is the keyboard. A keyboard allows the user to enter uppercase and lowercase letters, numbers, familiar symbols, and commands by typing them.

A typical keyboard contains the 26 letters of the alphabet, a space bar, common punctuation symbols, the digits from 0 to 9, the basic arithmetic operators, as well as a variety of command oriented keys like Delete, Backspace, Enter (sometimes called Return), and Shift and Control keys.

Most computer keyboards use the “qwerty” layout (where the alphabet is presented in the same way as with a traditional typewriter); however, other types of keyboards with a friendlier (more ergonomic) layout are obtainable.

When you press a key on your keyboard, a signal is sent from the keyboard to the computer along a wire that runs from the keyboard and plugs into the computer (typically through a USB port). There are also keyboards that can communicate with the computer using a wireless connection. Each key on the keyboard, when pressed, causes a different signal to be sent out, so that each key can be uniquely identified.

When the computer receives a signal from the keyboard, the signal is translated into the character that corresponds to the key that was pressed on the keyboard. At this point, the character is output to the computer’s display screen so that the user can see what they have just typed.

Keyboards are used wherever the user enters textual or numerical data. Together, the keyboard and mouse are the most frequently used input tools for personal computers.

The Mouse

A mouse is used by moving it over a typically flat surface. When it is moved over a surface, information about how the mouse’s position has changed is sent through a wire to the computer. (Just as with keyboards, there are also wireless mice available.)

When the computer receives this information, the mouse movements are translated into coordinates that can be applied to the computer’s display screen. This information is used to output a small image (called a mouse pointer) to the display screen. The movements of the mouse pointer on the computer screen will mirror the movements made by the actual mouse when it is moved about on a surface.

A mouse can be used to point to various items that are displayed on a computer screen. In addition, a mouse will have buttons and sometimes wheels, which relay an action command to the computer. This allows the user to point to an item with the mouse, and then click a button to perform an action on that item. In this way, a mouse can be used to start and stop programs, view the contents of folders, explore user interface components, and access software functions.

The Scanner

The scanner is a very interesting input device. This type of device can be used to convert an ordinary photograph, drawing, or sheet of text into a digital image. Basically, the object that is to be scanned (drawing, document, or photo) is placed flat on the scanner, and then the scanner is activated by the user.

Once activated, the scanner's optical sensing device records image information from the object that is being scanned (i.e. a photo or a drawing). This image information is processed by the scanner and is then fed into the computer through the scanner's connecting wire to a computer port.

Once received by the computer, the scanned image can be viewed on the computer screen, printed, stored on the computer's hard disk drive, or modified by using the appropriate software. As with many other input/output devices, the computer may require special software to communicate with the scanner.

In addition, the software packaged with some scanners will even help convert a scanned sheet of text into to an actual text file that can be edited and formatted, rather than just an image of the text.

Microphones

When it comes to audio data, computer speakers are probably the first device that springs to mind. While speakers make it possible to hear audio data that is output from a computer, another device, the microphone, allows a user to input audio data into a computer, just by speaking into it.

With a microphone, any sound (voice or other) that is made near the microphone can be stored on a computer. The microphone converts the mechanical energy of a physical sound wave into electrical signals that can be processed and stored on a computer. With VOIP technology, (Voice Over IP), a computer with a microphone and speakers can even be used much like a telephone. (There are also special VOIP phones available.)

A microphone will plug into a USB socket ion the computer. Most modern computers have a 'sound board' integrated into the motherboard or otherwise a separate sound board is inserted into connectors - the sound card is a device assists a computer process, input, and output sound.

When you speak into the microphone, your voice (sound data) can be recorded and stored on your hard drive. Microphones are often used with computers to add voice/narration to video clips, as narration or instructions, and to work with software applications that recognize speech, i.e. Microsoft Word.

Other Devices

For completeness, there are a few other input peripherals that are should be mentioned.

Trackball

A trackball is a pointing device that works like a computer mouse in reverse. With a trackball, you rotate a ball with your hand (fingers/thumb) to direct a pointer on your computer screen.

Unlike a mouse, which must be moved over a flat surface, a trackball mechanism is stationary. The movement data is generated by rotating the trackball within its stationary frame.

This means that the trackball does not require as much space as a mouse, and moreover, it will work on pretty much any type of surface.

Joystick

A joystick is another type of input device that feeds directional (and other information) into your computer.

The traditional joystick design consists of a bottom with a vertical hand grip that extends above it. When the hand grip is moved (tilted) in a particular direction, data is sent to the computer indicating this.

Joysticks are often used to play video games, or as a control for flight simulator software.

In addition to feeding directional information to the computer, a joystick will usually have one or more buttons to input other signals as well.

Touchpad

A touchpad is another type of input device that is used to help users navigate their computer screen.

Most often, touch pads are found on laptop computers, as they are relatively flat and only a few square inches in area. They are used to detect and translate the motions of a user's finger over the pad into corresponding movements of a pointer on the laptop screen.

Touch pads are used as mouse substitutes in laptops because they have a small profile that can be integrated with a laptop's compact size.

Computer Stylus/Pen

A stylus pen is another pointing device that is used to interact with a computer in a way that is like the use of a mouse or a touch pad.

The main difference is that the pen/stylus is used to "touch" or point directly to the object that you are interacting with (where it appears on your screen).

This means that that the pen is a direct or absolute pointing device, as compared to a relative pointing device like the mouse.

Pens or styluses are most often used with tablet PCs, graphics tablets (a kind of computer drawing pad), and with PDAs.

Digital Cameras

A digital camera can be used to capture digital images, which can then be stored on a computer for editing or display.

A digital camera is like a traditional camera in that it focuses light through a lens onto a specific area within the camera. In a traditional film camera, the light is focused through the lens onto a chemically treated film. The chemicals on the film react to the light and form (record) an image.

In a digital camera, light is focused onto sensors that convert the light energy into signals. These signals can be recorded as discrete values representing red, green, and blue color components. By recording color values for many small areas (pixels), a digital image is formed. These digital images can be stored on removable memory cards in the camera, or they can be fed into your PC (often via a USB connecter).

Many digital cameras can record videos and audio data as well as still pictures.

Lesson 2.3: Output Devices

In this lesson, you will learn about such common output devices as computer monitors, touch screens, printers, plotters, and speakers.

Monitors

Computer monitors or screens are probably the most essential of all the output devices. Modern monitors can provide a full color visual display of whatever the user is currently doing. When a user types something at the keyboard, it appears on the monitor. If a user clicks their mouse to open a file, the file contents are displayed on the monitor.

For most personal computer users, human computer interaction is made possible through the combination of keyboard and mouse as input devices, coupled with a monitor as an output device for visual data.

Monitors come in different shapes and sizes. Most modern monitors are LCD monitors (liquid crystal display). In a color LCD monitor, red, green and blue light are used to create a color display. The colour display is formed by liquid crystal molecules trapped between two plate-like surfaces which are realigned by passing electrical energy through them. When light is shone through the plates and liquid crystal molecules, it can be blocked or allowed to pass through, depending on the alignment of the molecules. This light is then filtered to produce combinations of red, green and blue.

All the colors that viewed on a computer screen are made from combinations of red, green, and blue light. Each computer screen is broken up into tiny discrete parts called pixels. Each pixel (picture element) on a screen can be made to emit light as described above. Anything you view on your computer screen is just a combination of specific pixels emitting different colors of light.

An important aspect of any monitor is its resolution. Resolution is a measure of how many picture elements (pixels) the viewing screen will be divided into. To some extent, if the screen resolution of the monitor is higher, the number of pixels used will be high and the display quality will be improved. (There are several other factors besides resolution that can affect display quality as well.) Modern screens will support multiple screen resolution settings.

The Printer

A computer printer is an output device used to produce a permanent, printed (paper) copy of an image, text document, or other output file from your computer. A file is selected on a computer (e.g. an image file or word processing document) and then a print command is executed. When the user performs a print command, the computer sends the data from the selected file through a wire to the printer for output.

What the printed copy of the file looks like can depend on what software program is being used to create or modify the file on the computer. Furthermore, different printers can produce hard copies of files with different quality levels and print speeds.

The main differences between various computer printers involve the mechanisms by which the printer produces its hard copy.

Probably the two most common types of printers in use today are laser printers and inkjet printers, both of which can produce good quality printouts. Laser printers use a toning compound like photocopiers, while inkjet printers use a very thin and precise spray of ink.

In addition, most current generation printers can produce full color printouts as well as printouts in black and white. For example, a color inkjet printer will use combinations of cyan, magenta, yellow, and black ink to produce color printouts.

Touch Screens

Touch screens are an interesting device in that they allow a user to input data, while at the same time outputting a visual display. Because of the input/output ability of touch screens, they are ideal for situations where you may not want a full keyboard or mouse as the primary input devices.

There are several ways that touch screens can be implemented, all revolving around the mechanism by which the point on the screen that is touched is determined.

With resistive touch screens, a change in electrical current, caused by the electrical resistance created when the screen is touched, is used to determine the location of the touch.

In capacitive systems, the electrical field of the touch screen is altered by the electrical capacitance of the object that touches it, like a finger or a conductive stylus held in the hand.

Other systems rely on such things as mechanical strain in the screen or optical sensors located on the screen's perimeter to pinpoint the location that an object touches the screen. Whatever the choice of technology, once the location of the touch is determined; the information can be input into the computer to interact with the object displayed on the screen at the location of the touch.

Speakers

Monitors and printers are probably the most obvious output devices; however, there is still another important output device that should be mentioned: speakers.

Computers can output data to a printer for a permanent hard copy and they can output video data to a screen so the user can see and interact with files and software programs. Computers can also generate, process, and output audio data which requires speakers. Note that nearly all modern computers come with basic internal speaker though many people seek an external speaker setup for higher sound quality.

Computer speakers are typically connected to a computer by plugging into the computer through a USB connector. Often computer speakers will come with their own controls for adjusting the audio tone (base, treble) and volume. Users can also control the volume and tone of the audio output by using software programs on the PC.

Plotters

A plotter is a type of output device that is very similar to a printer. The main difference is that plotters use a pen like device to draw continuous lines on the plotter paper, corresponding to positional information that is output from the computer.

Printers, on the other hand, produce image printouts based on a rectangular grid of pixels that taken together to form a larger image. Also, printers have a relatively narrow spectrum of paper sizes that they can accommodate, as they are often primarily used for printing text.

As a result, plotters can create a more finely resolved technically precise image or drawing than a typical printer. They can also create a much larger drawing because they are not as restricted as a printer in terms of paper size. As a drawback, though, they are also more expensive than printers. Today, plotters are most often seen in scientific, engineering, and architectural settings where very precise (and sometimes large) drawings or schematics are required.

By using multiple pens, each having its own color, a plotter can create multi-color drawings.

Lesson 2.4: Secondary Storage Devices

Storage devices are very important in a computer system. One of the main uses for computers is for the storage and retrieval of information.

In this lesson, we will take a closer look at computer storage devices, including internal and external hard drives, USB /flash drives, as well as tape drives and data cartridges.

Common Terms

A secondary storage device is a piece of hardware that is used to store data or programs that must be preserved if the computer's main memory fails or is shut down (powered off). This means that secondary storage devices must be able to retain data for some time even if there is no electrical power to the device.

Secondary storage devices all require some sort of medium to store data on. In a hard disk, data is stored on rotating disk platters. In a tape drive, data is stored on a long roll of magnetic tape. In a USB flash drive, data is stored in millions of tiny memory cells on a chip. Note that the use of floppy drives, zip drives and other traditional forms of data storage does not occur as these have been rapidly overtaken by new technologies which offer far greater storage capacity.

As an introduction to secondary storage devices, the following table lists some important terms that are used frequently used in the context of these devices.

Formatting	Formatting refers to the preparation of a storage device for use. Formatting often consists of the creation of an indexing system that allows the data items written on the storage medium to be organized and accessed in a consistent, efficient manner.

	The term formatting is most often used when speaking of disk drive devices (floppy disks, hard disks, and CD or DVD drives). In terms of hard disks, the formatting can vary depending on what file system the chosen operating system uses (such as FAT or NTFS).
Reading	The transfer of data from the storage medium into the computer's main memory.
Writing	The process of transferring data to the secondary storage medium.
Seeking	When speaking of hard disks and optical drives, seeking is the process of aligning the device's read write/heads over the correct part of the storage medium in preparation to read or write data.
Direct Access	This means that any part of the storage medium can be accessed directly. That is, the device can jump or move directly to the location of the data, as opposed to sequential access.
Sequential Access	The storage medium must be traversed in a sequential way to access the data. This means that some data can be accessed much more quickly than other data, depending on its location in the medium with respect to the reading mechanism.

Hard Disk Drives

A hard disk drive is a device that stores data on a stack of rotating magnetized platters. When the hard drive is in use, the platters rotate at a constant angular velocity (CAV). This means that when reading or writing data, the speed of the magnetic surface under the read write heads is constant.

To write data, information is sent from the computer to the read/write heads. These read/write heads, once they are in position, hover with only a tiny distance between them and the rapidly rotating surfaces of the platters. The information is then written by the write heads as electromagnetic signals on the surface of the rotating platter.

Because the read write heads are so close to the platters, and the platters rotate at a high rate of speed, hard disk drives can crash and fail if dust or dirt lodges between the read/write heads and the platter, or if the read/write heads contact the platters while they are rotating.
Hard disk drives have very large storage capacities; today, they are often in the hundreds of gigabytes range, and they can even be as large as 1 terabyte (roughly a trillion bytes) in capacity.

To read data from a hard disk, the file indexes created during formatting (tables that contain the locations of the data files on the hard drive) are checked. Once the location on the drive is known, the read/write heads seek (travel across the radius of the disk) until they are in the correct position. When the platters rotate under the read/write heads, the electrical signals are read from the disk and are interpreted as logical ones and zeros (data).

With a hard drive, a typical disk access requires two mechanical operations. First, the heads must perform a seek operation to get into position, and then the disks must rotate under the heads for

the data to be written or read. This introduces time lag into a hard disk data access, in the form of a seek time and a rotational latency. This makes hard disk drives much slower than main memory (RAM).

Hard disks provide reasonable access times for a secondary storage device, with data being accessed in times on the order of milliseconds (thousandths of a second). Hard disks provide direct access to the data that is stored on them. That is, the reading/writing mechanism can access any addressable part of the storage medium by moving directly there; the entire medium does not have to be traversed.

Basically, hard disk drives come in two major types: external and internal. Internal drives are housed inside the computer's box or tower, while external drives are enclosed in their own housing outside of the computer. An internal drive will generally offer data higher transfer rates and will be powered by a direct internal connection to the computer's power supply. An external hard drive will generally connect to the computer through a USB port.

One advantage of external hard drives is their portability, which allows the user to store or backup data from their computer onto the drive, take the drive to another computer system, connect it through a USB port, and then access the stored data.

USB Flash Drives

Currently, USB flash drives are arguably the device of choice for portable secondary storage. USB flash drives can be quite small or a rectangular shaped box varying in size. The unit will have a USB interface (connector).

Inside the USB's housing, chips containing a type of EEPROM (Erasable Electronically Programmable Read Only Memory) are used to store data. These memory chips provide nonvolatile storage like traditional ROM memory, but unlike traditional ROM, the stored data or programs can be erased with relative ease, and new data or programs can be written in their place.

To use the USB flash drive, a user just plugs it into a USB port, at which point it can be accessed through the computers operating system interface.

USB flash drives have storage capacities ranging from tens of megabytes to terabytes. Also, flash drives can provide read access times that are faster than hard disk drives. Because the flash drive has no moving parts, there is no seek time or rotational latency to account for as there is in hard drives. However, current hard disk drives still far exceed USB flash drives in terms of storage capacity however this is advantage is rapidly being eroded as technology improves. A USB flash drive can provide direct access to the stored data.

Advantages of USB flash drives are that they are durable and are relatively inexpensive. This combination of durability, speed, price and capacity are making USB flash drives the dominant external data storage devices.

Data Cartridges and Tape Drives

Tape drives are mechanisms that are in many ways much like a traditional tape recorder or video cassette recorder. Basically, a long spooled magnetized tape is passed under read/write heads to record or read data.

Tape drives provide sequential access to the stored data. This means that if the data that you want is positioned in the middle or end of the magnetic tape, and the tape must be rewound to the beginning, you must traverse the length of the tape until you get to the part that contains the data you are after. Because of this, they are much slower than direct access storage devices like hard disks and USB flash drives. However, the data that is stored on a tape can be preserved for a long time, and in addition, magnetic tapes are a relatively inexpensive storage medium. For these reasons, even to this day the servers in many companies, universities and other institutions still use tape drives to back up and archive their data.

Tape drives will typically use some type of data cartridge for storage of the actual magnetic tape medium. These tape cartridges usually consist of protective plastic housings around the spooled magnetic tape. The cartridges can be removed from the tape drives and archived for extended periods.

While tape drives may not provide fast direct access to data, they do provide a relatively inexpensive means of archiving very large quantities of data for extended periods of time.

CDs and DVDs

CDs (Compact Disks) and DVDs (Digital Versatile Disks) are two forms of optical storage media. Unlike most other secondary storage media, CDs and DVDs do not retain data in the form of electromagnetic signals on a magnetic surface. Instead, CDs and DVDs store data as physical features (called pits and lands) on the surface of a slim plastic disk.

When a sequence of pits and lands on a disk rotates under a beam of light, the light is reflected or scattered depending on what physical feature of the disk it is incident on (pit or land). These different light responses can be interpreted as binary data (a logical one or zero).

Unlike hard disk drives, CD and DVD disks can be loaded or removed from their drives for portability, or archival purposes. These inexpensive disks are quite durable, and data can be safely preserved on them for a long period of time. This makes CDs and DVDs a great choice for publishing music, software, and video.

CDs were originally designed as a publishing medium for music, so CD drive speed is measured relative to the standard established for audio CDs. A 1x speed CD drive will work at the same speed as a music CD, but a 4 x CD drive will work at four times the speed of an audio CD. Today, it is not uncommon to see CD drives with speeds exceeding 12x audio CD speed.

CDs and DVDs come in many varieties. A CD-ROM disk (Compact Disk-Read Only Memory) is a disk that data can only be read from. A CD-R disk is a disk (blank at first) that can be written to once

(though not necessarily all at one time), and then read many times. A CD-RW disk can be written to, erased, and rewritten repeatedly.

DVDs can be purchased in similar formats, with DVD-ROMs being used primarily for publishing of video, music, and software. There are also writable DVD-R and DVD +R disks, as well as rewritable DVD+RW disks.

The main difference between DVD disks and CD disks is storage capacity. A CD can typically hold between 600 and 700 megabytes of data, whereas a DVD can hold 4 gigabytes to 8+ gigabytes of data (depending on the type of DVD disk). Most computer DVD drives can also read CD disks.

A CD or DVD drive is typically located in the computer's box or tower with the front-loading mechanism accessible. A power cable will connect to the computer's power supply and a data cable will connect to the computer's motherboard.

Access times for CD and DVD drives are significantly higher than for hard disks because the optical drives must find the correct rotational speed for the disk depending on what part of the disk is being accessed. A hard disk, on the other hand, always uses a constant angular velocity (meaning a constant speed of rotation).

CDs and DVDs are mostly overlooked for data storage in today's computing world due to the relatively low capacity of the medium. Optical players are generally not included in the newer models of computers.

Storage Comparisons

To help you understand the wide variety of computer storage types, the following table provides information on a few of the more commonly used memory/storage devices, in the context of speed, capacity, and relative expense.

Storage/Memory type	Access Speed	Storage capacity	Expense
SRAM (cache memory) Volatile Not portable	Extremely fast (nanoseconds) Provides random access	Relatively low capacity (often only a few megabytes)	Price per megabyte is the highest
RAM (main memory) Volatile Not portable	Very fast (measured in nanoseconds, but is slower than SRAM) Provides random access	Fairly high capacity (in most PCs, hundreds of megabytes to 1-2+ gigabytes)	Price per megabyte is quite low, much more affordable than SRAM
Hard Disk	Moderately fast, (quite slow when	Very high capacity; hard disks with	Price per megabyte of capacity is very

Nonvolatile Not portable	compared to RAM) Access times measured in milliseconds Provides direct access	hundreds of gigabytes of storage are now common	low, works out to be less than the cost of RAM
USB Flash Drive Nonvolatile Portable	Quite fast (faster than a hard disk, but slower than RAM) Provides direct access	Good capacity; USB flash drives now offer capacities measuring in terabytes	Price per megabyte of storage rivals that of hard disk drives
CD Drives Nonvolatile and portable Portability, durability, and affordability make it a good data publishing medium	Slower than a hard disk. Best used as a publishing medium or for archival storage/backup Provides direct access	Moderate capacity, with each disk holding about 600 - 700 megabytes	Price per megabyte of storage is quite low
DVD Drives Nonvolatile and portable Great medium for publishing and/or archiving all kinds of digital information	Slower than hard disk Best used as a publishing/archival medium Provides direct access	Good capacity (4+ gigabytes)	Price per megabyte of storage is quite low

In general, the price of storage or memory will decease with access speed and memory density. (Fast, low capacity memory is more expensive than slower, high capacity storage.)

SECTION 3: SOFTWARE

In this section you will learn about:

- Operating systems
- Application software
- Software versions
- Software updates
- GUIs
- What an operating system does
- Some common operating systems
- What software applications do
- Some common applications
- Software analysis, design, programming, and testing
- Shareware
- Freeware
- EULAs
- Copyright
- Data protection legislation
- What to be aware of when downloading computer files and programs
- What to be aware of when sharing computer files and programs

Lesson 3.1: The Basics

Modern computers can help us with many tasks at home and at work. Computers form a key part of the global information/communication infrastructure and are now essential tools in almost all sectors, including health, transportation, communications, education, science, engineering and business.

It is important to remember that the power of computers comes from their programmability. Computers are tools that can be programmed to do a wide variety of tasks, allowing one computer to perform a variety of different functions.

Software is a term that collectively refers to the programs (instructions) that provide useful functionality to a computer system. Software is built from organized sequences of instructions that are interpreted by the computer's CPU to tell the computer what to do.

Software can be differentiated form hard wired instructions (ROM) in the sense that software can be loaded into a computer's memory and run, and then removed from memory so other software can take its place.

Because multiple software programs can be stored in RAM at the same time, and because a CPU can quickly switch from executing the instructions of one program, to executing the instructions of another, modern computers can run multiple programs at the same time. When multiple programs are running on one computer, each program gets its turn to access the CPU based on the management of the computer's operating system.

In this lesson, you will start to learn about software, the other main facet of computers. You will learn what an operating system is, what applications are, what versions mean, and what software updates are. Finally, you will also be introduced to the very important topic of GUIs (Graphical User Interface).

What is an Operating System?

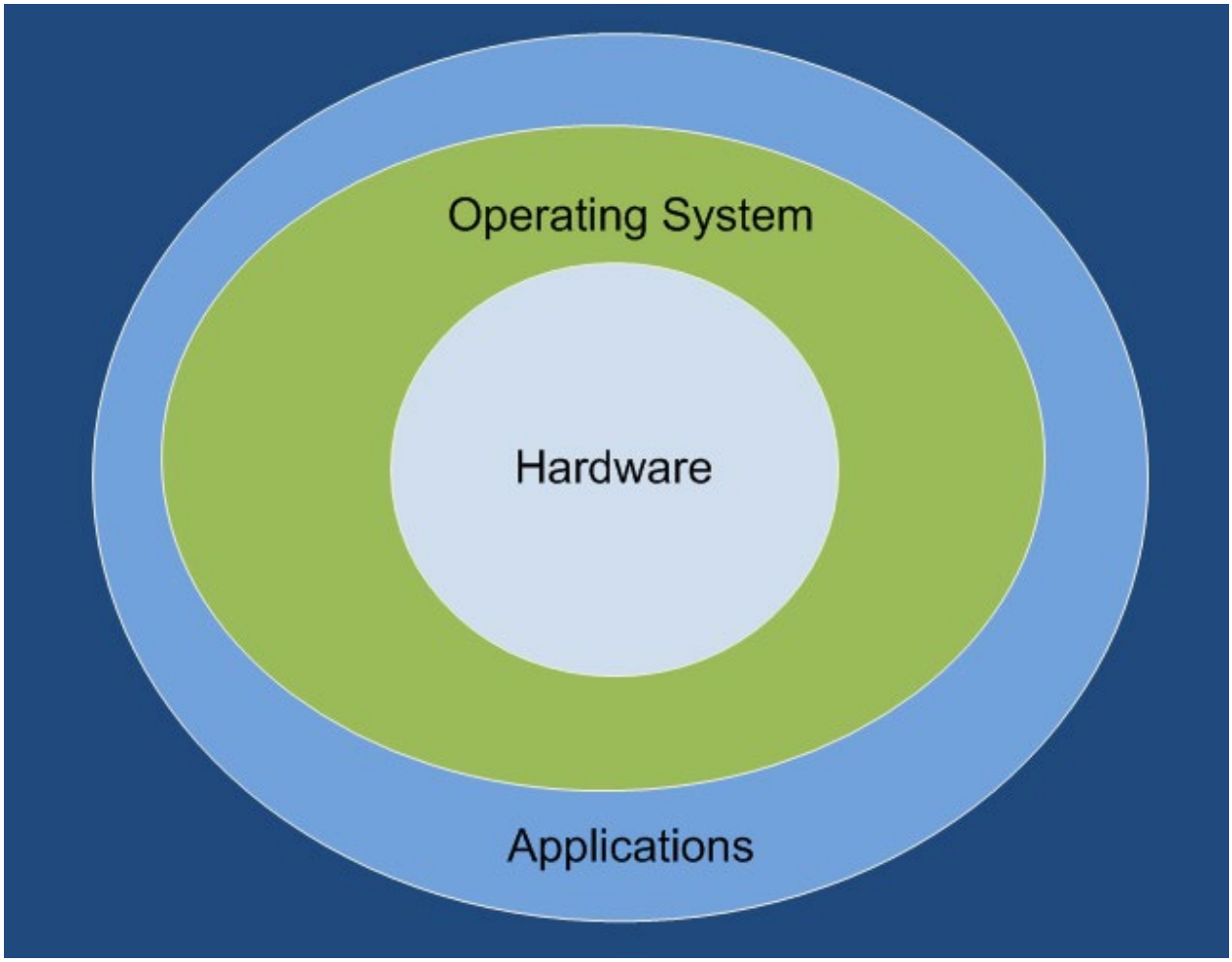
An operating system is arguably the most important software program that will run on a computer. An operating system provides an environment in which other software programs can run, while also providing functionality that allows users to interact with the computer.

The tasks that an operating system performs can be thought of as services. Some services are provided for users of the computer system, while other services are provided for the computer system itself.

Operating systems can vary to a great degree, depending on their intended purpose. Some operating systems are designed for multiprogramming environments or server systems, where multiple users log in and use system resources at the same time. Other operating systems are designed to provide a very user-friendly environment with easy access to system files, application programs, and functions. Whatever the case, all operating systems provide users with controlled access to the data files, software programs, and hardware that make up the computer system.

More specifically, the operating system is that program that is always running, whenever the computer is turned on. Other application programs may be arbitrarily started or stopped, but the operating system is always running until the computer is shut down.

You can think of an operating system as the software that is closest to the computer's hardware. The application programs that run on the computer (and the users who use these application programs) access the computer hardware either through, or with the help of, the operating system.



Often, operating systems are very large and complex pieces of software, with many thousands (or even millions) of lines of programming code. An operating system also has the important job of managing how the CPU accesses the instructions in different programs, when multiple programs are running at the same time.

What is an Application?

An application is a software program that is designed to perform a specific task or set of related tasks, e.g. there are applications that are designed to manage or interpret data (spreadsheets and databases) while some other applications are designed to help users create and modify text documents or drawings (such as word processors and vector graphics programs).

There are so many applications in use today that it can be difficult to list them all, but it is possible to differentiate between application software and the underlying systems software.

The job of the operating system is to create a working (operable) environment and at the same time manage and coordinate the resources of the computer system. An application, on the other hand, employs the system resources to accomplish a specific user-oriented task or set of related tasks. Furthermore, an operating system can function without any running applications, but applications cannot function without an underlying operating system.

Some software applications can be very large and complex, depending on their purpose, and they may require substantial time to be mastered by a user. Word processors, spreadsheet programs, and Web browsers are examples of common software applications that are in widespread use in businesses and in homes.

What do Versions Mean?

History shows that computer technology has evolved very quickly over time. Sometimes, after a software program has been created, new developments in computer performance, hardware and software design can make the software outdated or even obsolete.

It is also possible for software development companies to add additional functionality to their applications to stay even with competitors, to fix bugs (errors) in the program or to accommodate requests and feedback from users or clients. For whatever reason, software applications will often evolve over time as do the other aspects of computer technology. This means that a given software program can be available in many different versions.

Software versions are indicated by a numbering system involving three or more digits separated by periods (as in a decimal number like version 2.4.0). Sometimes, the digits are combined with letters, as in version 2.5b. In general, a higher version number for the software will indicate a more recent version.

Often when a software program uses the dotted number versioning system (e.g. version 1.2.0), the first number indicates the major release of the software (the digit one, in this example). If there is a relatively minor change to the version, the second number will be incremented. In this example, a minor change to the software might result in a number like version 1.3.0. The third number might be changed to reflect a small revision like the correction of an error or program fault. In the example, if an error was found in the program and then the error was fixed and the program re-released, the version number might be version 1.3.1.

Finally, if there has been a major change in the program's functionality or appearance, the first number in the dotted number notation could be changed (for example, from version 1.3.1 to version 2.0).

What are Updates?

It can be a very difficult (if not impossible) task for program developers to create large applications or operating systems that have no faults, flaws, security concerns, or other issues. It is often the case that new versions of a software program will evolve due to changes in technology and competitive pressures in the marketplace.

Sometimes, an older version of a software program can be converted to a newer version by installing a software update. Updates allow the user to upgrade his or her existing software to the latest version (or at least a more recent version) without having to install the new version from scratch.

Often, when a large software program (like an operating system) is released, there are several updates that follow as various security vulnerabilities, bugs or incompatibility issues with the software become known. These software updates are available over the Internet on servers provided by the company that make the software. These updates can be downloaded to the user's computer, and then installed to fix, improve, or otherwise update the given software program.

What is a GUI?

In the early days of computers, human-computer interaction took place on the command line. A user would type a specific command on his or her keyboard to execute some corresponding task on the computer system.

In these situations, a typical user had to memorize commands and command argument rules to interact with the system on even an elementary level. To be a truly advanced user, in-depth knowledge of the computer's file system, operating system commands and even some programming knowledge might be required.

Today most computer users interact with their operating systems and applications using GUIs.

A GUI (pronounced gooey) stands for Graphical User Interface, and it is the part of the application or operating system software that provides the user with access to the software's functions. A GUI consists of those objects on your computer screen that you interact with by dragging or clicking your mouse or by entering data with your keyboard.

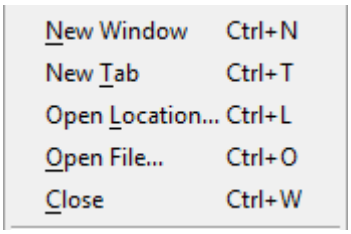
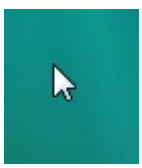
If you break the phrase "Graphical User Interface" down, the term Graphical refers to the graphics or images that appear on the screen representing underlying object in the system. The term User refers to the person that is interacting to the computer and the term Interface refers to a mechanism that facilitates the interaction between the user and the underlying system functions.

Putting this together, a Graphical User Interface uses **graphical** symbols corresponding to system components and functions, to allow the **user** to interact (**interface**) with these same system components and functions.

Because a graphical user interface relies on graphical symbols to represent underlying functionality and resources, the user can interact with the computer largely by using a simple pointing device (like a mouse). This allows the user to view data, devices, and applications on the computer as physical or tangible objects that can be easily opened, started, stopped, and removed, etc.

GUIs are used as an operating system interface and as an interface to the commands and functions within a software application. Different operating systems and applications can all have different GUI designs, but here are a few components that are consistent across most GUIs

The following table describes some commonly used GUI components.

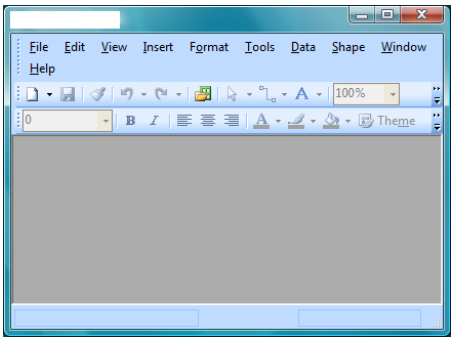
Icon		An icon is a symbol in the form of a small graphical image that corresponds to a device, a folder, a file, a program, or some other component of the computer system. By interacting with the icon, the user can manipulate the underling resource that the icon corresponds to. Icons are frequently used in operating system GUIs to allow users to access the computer's disk drives, files, applications, and more.
Menu		A menu is a panel of options that appears when a user clicks on a given menu heading or symbol. Menus are often used as a means of providing access to the functions and commands in a software application. To perform a command, the user clicks the corresponding menu item.
Button		A button is a small graphical item that gives the visual effect of being pressed when you click on it (using a mouse or another pointing device). Buttons are frequently used to access commands and functions in software applications. In a typical software application, most of the functionality of the application can be accessed through groups of buttons and/or menus.
Pointer		A pointer is a small graphic (often in the shape of an arrowhead) that is used in GUIs to show the current location where mouse commands (clicks or dragging) will be implemented.

The position of the mouse pointer will change on the screen corresponding to how the mouse is moved by the user.

Cursor

A cursor is usually a simple vertical line, underscore, or other symbol that flashes to indicate the location on the screen where any typed data will be entered.

Window



A window is a square or rectangular bounded area that encloses GUI components, data, and/or the working area associated with a given application or system component.

Windows of one form or another are used to access most programs and data on computers that use GUIs.

Typically, windows can be minimized, restored, and resized to accommodate the user's preferences for how the items on the display screen should be arranged.

There are many other GUI components available beyond what is shown above. A typical GUI will often include items like checkboxes, radio buttons, text fields, scroll bars, dialogue boxes, tool tips, and list boxes.

Lesson 3.2: Operating Systems and Applications

In this lesson you will learn what an operating system does and what a software application does. You will also be introduced to some of the most common operating systems and applications in use on today's personal computers.

What Does an Operating System Do?

An operating system provides an environment for application programs to run in and provides access to the underlying hardware devices and file system to both the program applications and the users.

The many functions of an operating system can be thought of as services that are provided for the user and for the computer system itself. Some of the more important services that an operating system provides are summarized in the following table.

System Services	User Services
The operating system provides CPU scheduling and management services. These important services essentially decide what programs have access to the CPU, and in what order, and for how long.	The operating system provides the user with access to the underlying computer hardware. Generally, to access the functions of the systems hardware, users must go through the operating system.
The operating system provides memory management services. This involves sharing the available memory among the various programs that may be running and allotting additional memory to programs that may require it. Memory management involves controlling access to main memory (RAM) and the use of secondary (hard disk) storage as well. By providing a virtual memory environment, operating systems can combine physically separate RAM storage and hard disk storage into one large addressable memory space.	The operating system provides the user with an interface to access to stored data. Users can access data through the operating system by using an interface (often a GUI) to navigate through the directories and files in the underlying file system. The operating system also allows users to access stored files through an appropriate application (like a text editor or database).
The operating system provides error handling services. If a program causes an error during its execution (usually detected by hardware), this error is sent to the operating system to be dealt with. The program can be stopped by the operating system or allowed to continue depending on the type of error. In some cases a message about the error will be output to the user.	The operating system allows users to run applications. An operating system provides an environment (often a GUI) that allows users to start, stop, and run programs that have been installed on the system.
The operating system provides control over the input/output devices. In this regard, the operating system manages how programs access input and output devices and resolves conflicts that may arise.	The operating system can provide security services for the user. An operating system can help protect a user's programs and files from unauthorized access (often using accounts with username and password combinations).

Essentially, an operating system controls and manages the available resources (memory, disk storage, CPU access, input/output devices) based on input from both software programs and users. The operating system maintains and manages a disk file system for storage of programs and data, provides an operating environment in which programs can run, and provides an interface for the user to access programs, data files, and hardware.

What are Some Common Operating Systems?

There are different types of operating systems that have been designed for different types of computer systems. An operating system that manages resources on a mobile phone does not necessarily require the same design strategies as an operating system for a PC or for a mainframe computer.

Even for a single category of computer system, there can be multiple operating systems available. The following table offers a brief description of some of the more common operating systems.

Microsoft Windows	There are several versions of the Microsoft Windows operating system with the most current being Windows 10. Microsoft Windows operating systems are currently the most common operating systems on desktop personal computer systems. Windows operating systems feature a very user-friendly GUI and wide compatibility with many software applications and PC Games.
UNIX	UNIX is a very powerful and flexible operating system. It is often used on servers and in multi-user/programming environments. With UNIX, the user can access the underlying system resources through a command line shell. There are GUIs available for UNIX systems, but many UNIX users enjoy the power and flexibility of the UNIX command line tools. UNIX is often used for large servers in businesses, universities, and other institutions.
Linux	Linux is another operating system that is very closely related to UNIX. Linux comes in many different versions, and in most cases, it supports a GUI of some kind. However, just like UNIX, Linux users often prefer the power and flexibility afforded by the command line. Linux is becoming increasingly popular as an alternative to Windows operating systems for personal computers. Linux, like UNIX, can also be used on servers. There is not as many desktop software applications available for UNIX or Linux systems as there are for Windows.

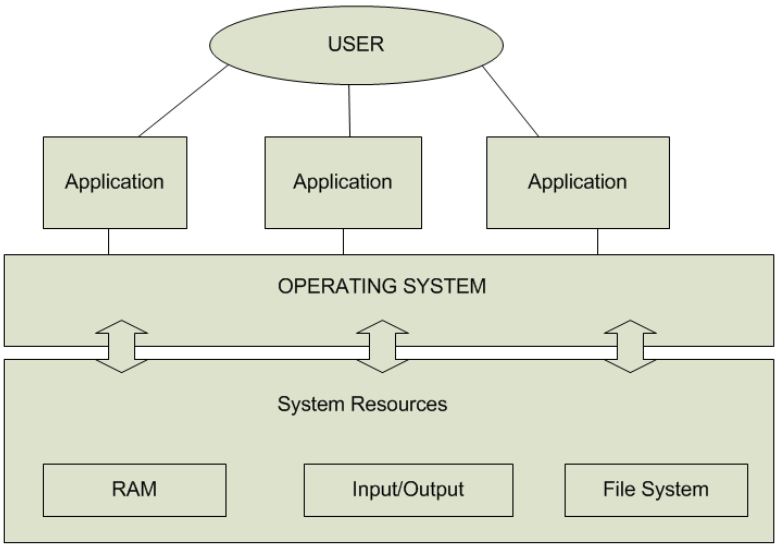
Apple OS	The Apple line of operating systems is designed for Apple's computers. Apple computers are personal computers that have a different architecture than most other PCs. Apple provides its own user-friendly GUI driven operating system for its own line of computers.
DOS (Disk Operating System)	The DOS operating system provides command line tools for managing files and programs. Early versions of the Windows operating systems were comprised of a Graphical User Interface that provided user friendly access to the underlying DOS system. Although more recent Windows operating systems do not run on DOS, they will still typically provide a DOS like command line environment.

What Does a Software Application Do?

While it is the job of the operating system to act as overall manager of the computer system, the job of a software application is to use the system resources to accomplish more specific, user-oriented tasks. An operating system typically performs a wide variety of tasks without having to be directed by the user. Software applications, on the other hand, often require extensive user interaction to accomplish the goal for which the given application has been designed.

A typical operating system will manage main memory, as well as maintaining and managing the file system without being explicitly told to do so. An application such as a word processor will not accomplish much of anything if the user does not interact with it.

Applications are the individual software tools that that help a user to accomplish specific goals, like creating a chart, editing a digital photo or preparing a report. Software applications access system resources (like files or hardware) in response to user actions and commands. The application uses these resources successfully because of the services provided by the operating system.



What are Some Common Applications?

Because there are so many tasks that people want to perform on their computers, there are a variety of applications available to accommodate them. The following table describes some of the more common application types in use today.

Word Processors	A word processor provides a user with options for creating, editing, and formatting text-based documents. With a word processor, users can choose from a wide variety of text fonts, sizes, and colors. Users also have fine control over page layout, margins, styles, alignment, and text effects. Modern word processors often include features like thesauruses and spelling dictionaries and can even automatically detect spelling mistakes and questionable grammar. The most common word processing application is Microsoft Word.
Spreadsheet Applications	Spreadsheet applications are numerically oriented software programs that help users store, organize, analyze, track, and report data. Typically, spreadsheet applications use small, indexed data cells to store and organize data, functions, and formulas. These cells are organized as columns and rows in a grid pattern. A cell is indexed according to what column and row it is found in. For example, cell A1 would be the cell in the first column of the first row. A function or formula in one cell can reference the contents of another cell by using the cell indexing system. This means that a complex formula or function can be defined in one cell that involves the values in many other cells. Spreadsheet programs are often used to create and calculate budgets, analyze numerical trends, and perform a host of other accounting and financial tasks. Microsoft Excel is a spreadsheet application.
CAD Applications	CAD stands for Computer Aided Design. CAD applications are frequently used in engineering, architectural, construction, and surveying/mapping settings. CAD software can be used to create architectural plans and engineering schematics, and for modeling three dimensional objects and surfaces. CAD software can be quite complex, and it is usually employed by users with knowledge in a specific (often technical) field. AutoCAD is an example of a CAD application that is frequently used in industry.

Photo/Image Editing Applications

Photo editing software allows users to modify, enhance, edit, and combine digital images. A well known digital image/photo editing application is Adobe Photoshop.

Software Development Applications (IDEs)

Software development applications (IDEs) provide environments that help programmers organize, create, compile, and test programming code. (IDE stands for Integrated Development Environment.)

IDEs will typically provide an editing window where programmers enter code. This editing window will help the programmer by color coding different statement types, language keywords, and variables. Furthermore, these development environments can help the programmer organize the components of a code project and provide tools that can assist in finding programming errors.

Microsoft's Visual Development Studio and the open-source program Eclipse are examples of IDEs.

Database Applications (DBMS)

Database applications or DBMS (Database Management Systems) are used to store, organize, and retrieve fields and records of data. This data is stored in an organized interrelated database often consisting of multiple tables of related information. Often, databases are queried (meaning data is requested) by other software programs.

Because databases can relate the stored information in a way that facilitates the query process, they allow programs and users to retrieve or filter data based on an extensive variety of criteria. Database applications (DBMS) are used by many organizations wherever there is a requirement for large amounts of stored but readily accessible information.

The term database can be used to refer (sometimes confusingly) to both the DBMS application and the actual stored data itself. Some common database applications in use today are Oracle, Mysql and Microsoft Access.

Web Browsers

Web Browsers are applications that provide users with access to the World Wide Web. A modern web browser is capable of understanding and rendering web pages authored with a variety of web development languages and technologies.

With the current popularity of the Internet, web browsers are probably one of the most frequently used applications that are found on desktop computers. Commonly used Web browsers are Microsoft's Edge, Mozilla's Firefox and Google's Chrome.

Other Applications

There are many other software application types in use today, including but not limited to:

- Video editing applications

- Security applications (firewalls and anti-virus programs)
- Computer animation programs
- Peer to peer (also known as P2P) file sharing programs
- Optical disk authoring (burning) programs
- Audio/music recording and editing software

Lesson 3.3: How is Software Built?

In this lesson, you will introduce to some of the aspects of a typical software development/engineering process, including the analysis stage, the design stage, the programming stage, and the testing stage.

Introduction

Most modern software applications are written in a high-level programming language. A programming language consists of a human-readable set of rules and specific keywords that can be used to precisely define the instructions used to make the computer do what is required. This set of human-readable instructions written in a high-level programming language is called the program's source code.

Because source code is human readable, it can typically be understood, edited, and modified by those people who are trained in the use of the language that the code was written in.

When a program is ready, the human readable source code is compiled by a special software program (or set of programs) called a compiler. Compiling is the process of converting the high-level source code into a machine language that can be understood (processed) by the CPU. Typically this machine language consists of sequences of ones and zeros (bits) that specify precise instructions or tasks that can be interpreted by the CPU's control unit.

To summarize, computer software can be thought of in two ways: as human readable source code and as a set of binary instructions (machine language) that the CPU can process.

The source code for a program is developed by programmers according to a specific software development process. Once it has been deemed that the software is ready, it is released (typically in the form of machine language) so users can install and run it on their computers.

To produce good quality software, programmers will generally follow a carefully designed plan or process. The art/science of developing software comprises a field of its own called software engineering. The following pages will provide you with a gentle introduction to some of the stages in a typical software engineering process.

Analysis

In the analysis stage, software developers try very hard to capture all the requirements of the software that they will build. A requirement is a task that the software must accomplish, or some specification for a feature that the software must provide to achieve its overall purpose.

Capturing requirements involves interviewing potential users and studying the systems that software will be implemented on as well as the environment that the systems will be operated in, all while keeping in mind the software's overall purpose.

At this stage, the requirements that are being captured can be organized into many requirement types. Nearly every requirement can be grouped into one of two general categories: functional requirements and nonfunctional requirements.

Functional requirements are tasks or features that must be performed for the software to correctly fulfill its overall purpose. For example, a functional requirement might describe how the software should act when a user executes a particular command, or how the software should behave when an error occurs.

Nonfunctional requirements are constraints or boundaries placed on the problem that the software must work within. For example, it may be a requirement that the software work with a particular type of operating system or on a particular hardware platform. It is also possible that the software must finish its tasks within a specific time.

Essentially, properly designed software should be able to complete its functional requirements within the boundaries or constraints imposed by its nonfunctional requirements.

Another important aspect of the analysis phase is to make sure that the requirements specified for the software are as complete as possible. Have user requirements, data requirements, interface requirements, and security requirements all been accounted for?

The analysis stage and the capturing of requirements are very important because it can be both costly and difficult to introduce missed requirements in the later stages of the development process.

Design Stage

Once the analysis stage is complete, and the requirements of the software have been specified, a high-level model of the software can begin to take shape in the design stage.

In the design stage, developers can discuss what strategies are the best for tackling the requirements of the software and begin to plan the different components and modules and supporting materials needed for the software to perform what is required of it.

In the design stage, the developers might ask questions like the following:

- What form should the user interface take to meet the expectations of the users?
- How should the output or results that the software produces be presented?
- How will the software get its input?
- What algorithms are necessary to process the data as required?
- Which program components should be responsible for which functions?
- How should the different modules or components of the program interact?
- What programming language is best for this application?

In general, as the details and requirements are worked out, the design stage will gradually move from very high-level design ideas (perhaps pencil and paper sketches of interrelated components) to more precise specifications for the various software components.

Also, once a general design has been agreed upon, prototypes can be built. (Prototypes are hollow mockups of certain aspects of the software that lack most of the actual function. A prototype may take the form of a rudimentary graphical interface, which can be tested with users to see if any requirements have been missed, or any new requirements can be discovered.)

Eventually, the design stage takes the development process to a point where there are precise specifications and definitions for the components required by the application, for what each component must do and for how the components will work together (interact) as a single application.

At this point, the development team can begin the actual programming or coding of the application.

Programming

In the programming stage, programmers start writing code to meet the software specifications developed in the design stage.

Often, if the software application is large and entails many components, different programmers or teams of programmers will be assigned to the different components of the application, e.g. one team may be assigned to the development of the user interface while other teams are assigned to the core functional components.

The programming team breaks the software application into smaller, more manageable chunks to make the process of building the software easier and more efficient. Programmers who are responsible for a given component need only concern themselves with how that component functions and how it must interact with other components.

During the programming stage, the programming teams will typically develop the software using the same IDE (Integrated Development Environment), naming conventions (e.g. for variable names), and the same family of code libraries. (Libraries consist of existing code that can be added to a program to provide functionality.) Keeping these elements consistent helps the programmers avoid errors and incompatibility issues between components.

After they write code, programmers will compile and run it to see if there are any obvious errors. If errors are found, the source code is reexamined to find the error. When the error is found, the source code is changed by the programmers to remove the error. Then the code is compiled and run again to see how it performs. This process is repeated until the code appears to work correctly.

Building software can be a challenging job, and there are many places where faults can be introduced into a software application. If, in the analysis stage, the full requirements are not captured, this missing functionality will be passed on through the design stage and into the programming stage.

Flaws in the design of the software may lead to inefficiency or incompatibility issues. Moreover, poorly designed software can be difficult to maintain or update as time goes on.

In the programming stage, there are a host of errors (bugs) that can be introduced, involving the program's logic and flow and the actual grammar and syntax of the programming language itself. For these reasons, testing is a very important part of the software development process.

Testing

Software applications are often designed and programmed as multiple functional components or units that will eventually come together to form the completed program. Breaking a large software project down into components not only helps make the programming tasks more manageable, but it also provides an opportunity for extensive and thorough testing.

Because the development team knows the specifications for each software component (the requirements that it must meet), each software component can be tested on an individual basis. One advantage that this strategy offers is that if errors are found on the component/unit level, you will have a smaller amount of code to examine to find the cause of the error. If the software was written and tested as one huge program, it might be much harder to find where the error is located. Also, building software in a modular approach allows you to redesign and replace a single component if it does not work, rather than trying to redesign an entire application.

The testing of individual software components each on their own is called unit testing. If full requirements of a single software component are known, then the individual component can be tested to see if it meets those requirements. Also, because the units are tested independently, some units can be tested even if other units are not yet finished or ready for testing.

After the unit testing is complete, and it is deemed that the individual components perform as expected, the components can then be tested to see how well they work together, in a process called integration testing.

Integration testing involves putting individual software components together and testing how they function and interact when they are combined to form a larger entity. There are different strategies that can be employed for integration testing such as the top down or bottom-up approach, or the big bang method.

In the top down or bottom-up approach, an individual component is tested with the components that it must interface with directly. When this round of testing is complete and errors are resolved, the component or components that were tested with the first component are in turn tested with additional components that they must interface with. Continuing this strategy, the components are integrated and tested in layers, until all the components have been integrated and the entire software application is being tested as a unit.

With the big bang approach, the components are all put together to form a larger application, and then everything is tested together as a whole all at once (hence the term "big bang").

In both the unit testing stage and the integration testing stage, white box testing and black box testing are two strategies that are commonly used.

White box testing involves the testing of software by those who know the inner workings (code structure) of it. This allows the testers to use their knowledge of the code to come up with scenarios that test every potential logic path in the unit being tested.

Black box testing is quite different. This testing is done without the knowledge of the inner workings of the software. In this form of testing, testers try different values in the expected range of inputs for the code and see if the expected results come out. They will also try different values that are outside of the range of expected inputs to see how the software reacts.

Once the software has passed integration testing, it will often undergo system testing to see how the complete integrated software package meets the requirements specified for it.

Good software development strategies are designed to keep the introduction of errors to a minimum. Although testing is certainly essential, it is not a good idea to try to test quality into a software product. Instead, quality should be infused into the software through a well conceived development process.

Lesson 3.4: Types of Software

There are many different software programs available to PC users. Some can be freely downloaded over the Internet while others can be purchased and downloaded. Some software, like an operating system, will often come pre-installed on a computer.

In this lesson, we will look at some of the different categories that software falls under, like shareware and freeware. We will also consider other aspects of acquiring software, such as EULAs (End User Licensing Agreements) and how to check the version and the product ID number of a software application.

What is Shareware?

Shareware is software that is available for users to try for free, usually by downloading and installing it from the Internet. With shareware, after a set time, the user will typically be required to pay a fee for the software if they find the application to their liking and would like to continue using it.

The payment of a fee for shareware is voluntary and in some cases individual users will not be prevented from using the software if they do not pay for it. In many cases, if the user does pay the fee and register the shareware product, they will be able to take advantage of improved or additional functionality in the program.

Unlike traditional commercial software, the copying and distribution of shareware is not always discouraged. The general thinking is that the more users who try the software, the more users will pay the requested fee for it. Shareware is often developed by individual programmers or sometimes small groups of programmers, so shareware programs may not be as large or comprehensive as applications produced by a major software company. Nevertheless, shareware programs can have a high standard of quality.

To summarize, the basic idea behind shareware is to give the user an opportunity to try the software before they buy the software.

What is Freeware?

Freeware is computer software that is freely available for an unlimited time and at no cost. Often, freeware is developed by community minded programmers who would like to see their software more widely distributed among other users.

Just because freeware is free does not mean that the user can do whatever he or she wants with it. Freeware can still have license agreements that discourage unauthorized distribution or modification of the software. Freeware is copyrighted so the creator of the freeware program remains in control of the future development of it.

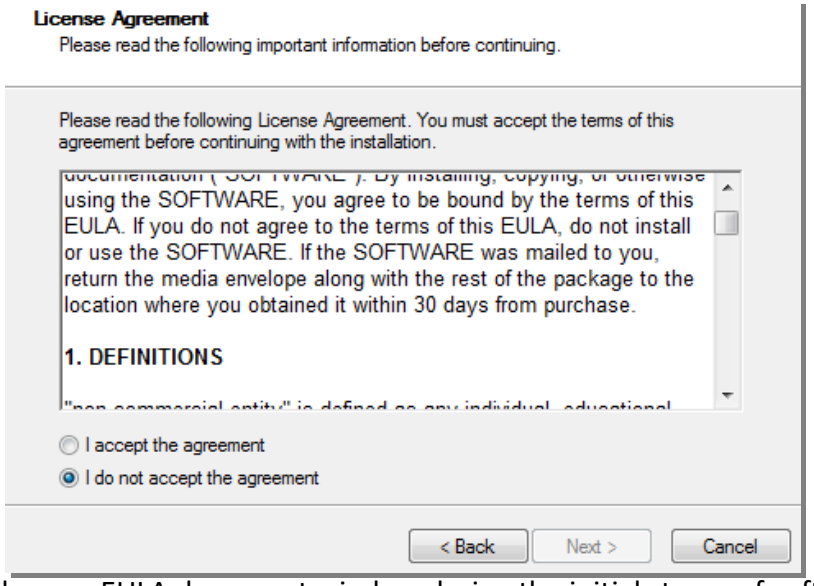
Like shareware, freeware is often distributed to users by making it available in a downloadable form over the Internet.

What is a EULA?

EULA stands for E**nd User License Agreement**. Most commercial proprietary software will include some form of user agreement specifying the conditions of the software's use.

Most EULAs are integrated with the installation process and appear near the beginning of the installation of the software in the form of an electronic document. Normally the document details the copy, distribution, and usage rights that the software producers want the users to adhere to.

If the user wants the software installation to continue, they will typically have to signify agreement to the terms of the document by clicking a checkbox, radio button, or an "I Agree" button. Once it has been indicated that the EULA has been agreed to, the installation of the software will continue.



The image above shows a EULA document window during the initial stages of software installation. Currently, debate over the enforceability of EULAs continues. However, they are in very widespread use whenever software is being sold or distributed.

How Do I Check the Software Version?

There may be times when you want to check the version number for a software application that you use. You can check the software version from within the application itself.

With the application started, you should search the user interface for an option that will provide information about the application. Often this option will be found on the program's Help menu and will be labeled with the word "About," followed by the name of the application.

When you click the About option, a window will appear that will typically display the name and version number of the software.

How Do I Check the Product ID Number?

Many software applications have a product ID number. This number can be used to identify the software, and in some instances, to determine if your copy of the software is a valid and licensed. In some cases, a valid product ID might be required for you to receive updates for the software.

The product ID can often be found in much the same way as the Version number. To start, look for a Help menu, and choose the "About" option if there is one.

At this point, a window will typically open that displays the version number and the product ID.

Lesson 3.5: Legal Issues

Software piracy is a major concern in today's software market. The very nature of software makes it particularly vulnerable to copying because of the way software is published, distributed and used. In this lesson, you will learn about copyright and how it applies to software and other digital media files. You will learn about data protection legislation, and what you should be aware of when downloading, using, or sharing software materials.

What is Copyright?

Copyright can be described as a set of rights that can control how a particular creation is used or copied. For example, if you write a book or story and it is copyrighted, there will be legal limitations imposed on how others can use your book or story. Only you will have the exclusive rights to copy and/or distribute your literary work as you see fit. The limitations on what others can do with your literary work are intended to ensure that you are benefiting from what you have created.

There are, however, certain materials that a copyright does not protect. For example, materials that lack originality and are just compilations of public data (like a phone book) cannot be copyright protected. A reprint of material that is already in the public domain cannot be protected by copyright.

Any material that is in the public domain (can be freely used by anyone) is not copyright protected. These materials include works that were created and published a long time ago (several decades) and most government documents.

Copyrighted software is treated much like a copyrighted literary work (even though there are many significant differences between software and a literary work). The unauthorized copying or redistribution of copyrighted software is forbidden and is deemed a copyright infringement. If a software user performs a copyright infringement on the exclusive rights granted to the creator of the software, the user may be held liable.

Some examples of possible copyright infringements are:

- Making unauthorized copies of the software
- Deliberately using an unauthorized copy of a software application
- Distributing unauthorized copies of the software (sharing with friends, coworkers, etc.)
- Purchasing a software application that is licensed for a single computer, and then using it on multiple computers

It is understood that a user must make at least one copy of a software application when they install the program to the computer's hard disk (a copy of the information on the publishing medium is made and saved on the computer's hard disk drive). Beyond that, a user should consult the EULA provided with the software for further details.

What is Data Protection Legislation?

Data protection legislation typically refers to government legislation that upholds the rights of citizens to have data privacy. With the advent of computer technology, tremendous amounts of private data are being stored on computers everywhere.

Some of the private information stored on computers may include:

- Personal information, like names, ages, and addresses
- Credit reports
- Phone numbers
- Credit card numbers
- Resumes
- Health records
- Various financial/banking data
- Criminal records
- Tax information

To keep this growing amount of stored data in check and to protect people's privacy, many countries have adopted data protection legislation that specifies guidelines for how this data is to be maintained. In Australia, the Privacy legislation specifies that data and data collection must be:

- Fairly and lawfully processed
- Processed for limited purposes
- Adequate, relevant, and not excessive
- Accurate
- Not kept longer than necessary
- Processed in accordance with the data subject's rights

- Secure
- Not transferred to countries without adequate protection

This legislation helps to enforce that organizations give people notice when they collect data, that the collected data is only used for the purposes stated, that the data will only be disclosed with the given person's consent, and that people will have the option of seeing what data an organization keeps on them.

How Does Copyright Apply to Software vs. Files?

There are many types of files that a computer user may have to work with. The files that are used on a computer will be created with software that is protected by copyright law. What does this mean for the files that a user creates?

A file that is created consisting of original content is the user's creation, even if the software used to create the file is copyrighted. A person can copy the content of such files as much as they wish and distribute them to whoever they like. There are files that should not be copied or distributed without the permission of the files' owners or creators.

Just as with software, computers make it very easy to duplicate music, video, and practically any other digital (computer) files without losing any quality. Peer to peer file sharing applications make it very easy to distribute these files over the Internet. Just as with software, the unauthorized copying, sale or distribution of copyrighted digital files may be copyright infringement.

Some common files (copyrighted materials) that are frequently copied and distributed without authorization are:

- Music files (MP3 files)
- Video files (DVD movies, MPEG/MP4 videos)
- Digital Images (JPEG, GIF and bitmap files)
- Software program files (.exe files)

What Should You Download?

There are vast amounts of files available for download over the Internet. The files may be available on Web sites, available through various peer to peer sharing programs or hosted on various file sharing servers.

When you download a file from an unknown source, you really have no idea what you are getting. The file may contain a virus that can damage other files on your computer and render your operating system inoperable. Another possibility is that the file you download may be copyrighted, and as such, you risk copyright infringement by downloading it and using it without authorization.

To avoid viruses, spyware, and other malicious code that can hide in downloadable programs, avoid downloading files from Web sites that you do not trust completely. For example, if you want to download and install an update for a Microsoft operating system, download it from the official Microsoft Web site.

When you download software to install on your computer, make sure you read and understand the EULA that comes with the software. If you are not comfortable with the terms of the agreement, you will have the option of canceling the installation.

Remember, if you download a software application from any source other than the software company's official site, there is a chance that you could be downloading an unauthorized copy. If you find a well-known software application available for download through peer to peer file sharing or any other source than the manufacturer's Web site, it is very likely unauthorized.

What Should You Be Aware of When Using Materials?

If you are using proprietary software applications or other files that are copyrighted, make sure that you are using them within the constraints of the accepted agreement. You should not try to reproduce or distribute these materials in a way that leads to copyright infringement. Unauthorized use of computer software (and other digital files) may deprive the developers of the rewards which they are entitled to.

Furthermore, using unauthorized computer software, or viewing and duplicating copyrighted music or video files (or any other copyrighted files for that matter) is, at the least, deemed unethical by many, and may even make you liable for copyright infringement.

What Should You Be Aware of When Sharing Materials?

When you share materials, you should realize that unless the materials have been created by you, or unless you have explicit permission from the material's actual creator to share them, you may be performing copyright infringement.

This is especially true when it comes to computer software and other digital files. When you share copies of an operating system or other software application, make sure you are doing so in accordance with the software's licensing agreements.

For example, if you are installing an operating system on multiple computers at your office, make sure that the number of installations does not exceed the number that is permitted by the terms of your licensing agreements. If you are taking copies of software home from work to install on your home computer, you may be infringing on software copyrights.

Similarly, if you take a copy of a software application that you have purchased, copy it to a USB, and give it to your friend to install on his or her computer, this may be considered pirated software.

Finally, if you download software, music files, or videos over the Internet to share with other users (perhaps with a peer to peer file sharing application), you are likely distributing unauthorized copies of copyrighted material.

Modern computer technology makes it very easy to copy and distribute all kinds of materials in many ways. Just because copying and sharing digital material is easy, it is not necessarily ethical or legal.

INDEX