

Hardware and Software

In thinking about *The Wizard of Oz*, you are thinking about information, not about something physical. You would not usually think about the actual material used to store the information (movie film or a DVD) nor about the device which uses the information (a movie projector or a DVD player). A DVD is physical, but the movie itself is intangible information.

When you speak of a movie, you usually mean the intangible movie, not a particular device that has a record of it. (Although human language is wonderfully flexible; if you were in a video store and someone asked you to hand them "The Wizard of Oz" you would probably hand them the correct DVD.)

The word *hardware* is used for physical devices such as TV sets, DVD players and computers. The word *software* is used for the information used with such devices: movies, music, novels, web pages, computer programs, and data.

When talking about computer systems, *hardware* means the physical parts of the computer. *Software* means the programs and data used with the physical computer.

QUESTION 2:

Imagine that you have a copy of the book *Tom Sawyer*. You tear it to pieces and push the shreds through your garbage disposal. (It was probably assigned reading.) Have you destroyed the novel *Tom Sawyer*?

Answer:

No. You have destroyed a particular physical book, but not the information that is the novel *Tom Sawyer*. (You will still have to write that book report.)

Components of a Computer System

Information is funny stuff. The information for the web page you are looking at is recorded on a hard disk inside a computer located in New Britain, Connecticut. The information was copied from that physical device to other physical devices perhaps hundreds of times before it got to your computer where your web browser is using it to display these words on your monitor. Clearly information is something different from the devices used to store and transmit it.

A computer system consists of both hardware and information stored on hardware. Information stored on computer hardware is often called *software*.

The **hardware** components of a computer system are the electronic and mechanical parts.

The **software** components of a computer system are the data and the computer programs.

[Click Here](#) to see the desktop system these notes were originally prepared on. (This was back when anything to do with computers had to be beige. Everything in this picture has since been replaced, except for the gargoyle.)

The major hardware components of a computer system are:

- Processor
- Main memory
- Secondary memory
- Input devices
- Output devices

For typical desktop computers, the processor, main memory, secondary memory, power supply, and supporting hardware are housed in a metal case. Many of the components are connected to the main circuit board of the computer, called the *motherboard*. The *power supply* supplies power for most of the components. Various input devices (such as the keyboard) and output devices (such as the monitor) are attached through connectors at the rear of the case.

QUESTION 3:

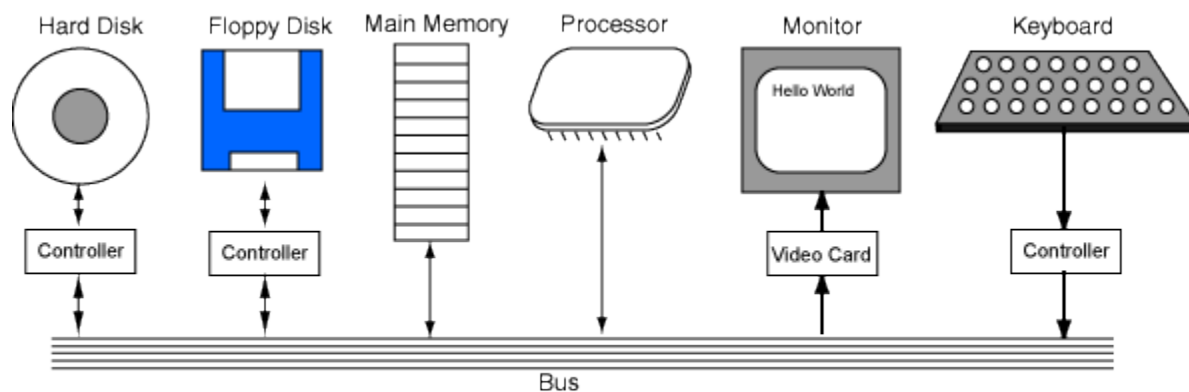
Is a *mouse* an input device or an output device?



Answer:

A mouse is an input device.

Hardware Components



Main Components of a Computer System

The terms *input* and *output* say if data flow into or out of the computer. The picture shows the major hardware components of a computer system. The arrows show the direction of data flow.

A **bus** is a group of wires on the main circuit board of the computer. It is a pathway for data flowing between components. Most devices are connected to the bus through a **controller** which coordinates the activities of the device with the bus.

The **processor** is an electronic device about a one inch square, covered in plastic. Inside the square is an even smaller square of silicon containing millions of tiny electrical parts. A modern processor may contain billions of transistors. It does the fundamental computing within the system, and directly or indirectly controls all the other components.

The processor is sometimes called the **Central Processing Unit** or **CPU**. A particular computer will have a particular type of processor, such as a Pentium processor or a SPARC processor.

QUESTION 4:

A modern processor chip (in 2014) may have up to 2.5 billion transistors.

How many transistors do you think the first commercial processor chip had in 1971? Click on a guess:

- 2.5 hundred transistors
- 2.5 thousand transistors
- 2.5 million transistors
- 2.5 billion transistors

Answer:

The first microprocessor chip, the Intel 4004, had 2300 transistors in 1971.

Memory

The processor performs all the fundamental computation of the computer system. Other components contribute to the computation by doing such things as storing data or moving data into and out of the processor. But the processor is where the fundamental action takes place.

A processor chip has relatively little memory. It has only enough memory to hold a few instructions of a program and the data they process. Complete programs and data sets are held in memory external to the processor. This memory is of two fundamental types: main memory, and secondary memory.

Main memory is sometimes called **volatile** because it loses its information when power is removed. Secondary memory is usually nonvolatile because it retains its information

when power is removed. (However, it needs power when information is stored into memory or retrieved from it.)

Main memory is sometimes called **main storage** and secondary memory is sometimes called **secondary storage** or **mass storage**.

- **Main memory:**
 - closely connected to the processor.
 - stored data are quickly and easily changed.
 - holds the programs and data that the processor is actively working with.
 - interacts with the processor millions of times per second.
 - needs constant electric power to keep its information.

- **Secondary memory:**
 - connected to main memory through the bus and a controller.
 - stored data are easily changed, but changes are slow compared to main memory.
 - used for long-term storage of programs and data.
 - before data and programs can be used, they must be copied from secondary memory into main memory.
 - does not need electric power to keep its information.

QUESTION 5:

Which is fastest? Main memory or secondary memory?

Answer:

Main memory.

Main Memory

Main memory is where programs and data are kept when the processor is actively using them. When programs and data become active, they are copied from secondary memory into main memory where the processor can interact with them. A copy remains in secondary memory.

Main memory is intimately connected to the processor, so moving instructions and data into and out of the processor is very fast.

Main memory is sometimes called **RAM**. RAM stands for **Random Access Memory**. "Random" means that the memory cells can be accessed in any order. However, properly speaking, "RAM" means the type of silicon chip used to implement main memory.

When people say that a computer has "512 megabytes of RAM" they are talking about how big its main memory is. One megabyte of memory is enough to hold approximately one million (10^6) characters of a word processing document. (There will be more about bytes and megabytes later on in these notes.)

Nothing permanent is kept in main memory. Sometimes data are placed in main memory for just a few seconds, only as long as they are needed.

QUESTION 6:

Imagine that you are using a computer to write a book report on *Tom Sawyer*.

- Where is the word processor program that you are running located?
- Where are the characters you have typed?

Answer:

- Where is the program you are running?
 - In main memory.
 - Running programs are always located in main memory.
 - A permanent copy will also be in secondary memory on the hard disk.
- Where are the characters you have typed?
 - In main memory.
 - When you "save" your document, the characters are copied to a file on the hard disk.

Secondary Memory

Secondary memory is where programs and data are kept on a long-term basis. Common secondary storage devices are the hard disk and optical disks.

- The hard disk has enormous storage capacity compared to main memory.
- The hard disk is usually contained inside the case of a computer.
- The hard disk is used for long-term storage of programs and data.
- Data and programs on the hard disk are organized into files.
- A **file** is a collection of data on the disk that has a name.

A hard disk might have a storage capacity of 500 gigabytes (room for about 500×10^9 characters). This is about 100 times the capacity of main memory. A hard disk is slow compared to main memory. If the disk were the only type of memory the computer system would slow down to a crawl. The reason for having two types of storage is this difference in speed and capacity.

Large blocks of data are copied from disk into main memory. The operation is slow, but lots of data is copied. Then the processor can quickly read and write small sections of that data in main memory. When it is done, a large block of data is written to disk.

Often, while the processor is computing with one block of data in main memory, the next block of data from disk is read into another section of main memory and made ready for the processor. One of the jobs of an operating system is to manage main storage and disks this way.

Primary memory	Secondary memory
Fast Expensive Low capacity works directly with the processor	Slow Cheap Large capacity Not connected directly to the processor

QUESTION 7:

Say that you have purchased the game *Doom IV* and have installed it on your computer. The game consists of

- A program (that controls the action of the game)
- Data (the images and other information)

After installation, where is each of these parts located in your computer system?

After you have installed the game, where are its parts located?

Answer:

- A program
 - On the hard disk
- Data
 - On the hard disk

When you play the game, its various parts are copied from the hard disk into main memory as needed.

Input and Output Devices

Input and output devices allow the computer system to interact with the outside world by moving data *into* and *out of* the system. An *input device* is used to bring data into the system. Some input devices are:

- Keyboard
- Mouse
- Microphone
- Bar code reader
- Graphics tablet

An *output device* is used to send data out of the system. Some output devices are:

- Monitor
- Printer
- Speaker

A network interface acts as both input and output. Data flows from the network into the computer, and out of the computer into the network.

QUESTION 8:

Do you think that data transfer from the network is slower or faster than from main memory?

Answer:

Data transfers from the network are much slower than from main memory and even slower than from hard disk.

I/O

Input/output devices are usually called **I/O** devices. They are directly connected to an electronic module attached to the motherboard called a **device controller**. For example, the speakers of a multimedia computer system are directly connected to a device controller called an audio card, which in turn is plugged into a bus on the motherboard.

With many recent computers, the functions of a device controller are integrated with the motherboard. Some motherboards have audio, graphics, and network controllers built in.

Sometimes secondary memory devices like the hard disk are called I/O devices (because they move data in and out of main memory). What counts as an I/O device depends on context. To a user, an I/O device is something outside of the computer case. To a programmer, anything outside of the processor and main memory is an I/O device. To an engineer working on the design of a processor everything outside of the processor is an I/O device.

QUESTION 9:

Is a game controller an input device or an output device?

Answer:

Input. (Although some game controllers have tactile feedback and so are output devices as well as input devices.)

Embedded Systems

A computer system that is part of a larger machine and which controls how that machine operates is an **embedded system**. Usually the processor constantly runs a single control program which is permanently kept in ROM (Read Only Memory).

ROM is used to make a section of main memory read-only. Main memory looks the same as before to the processor, except a section of it permanently contains the program the processor is running. This section of memory retains its data even when power is off.

A typical embedded system is a cell phone. This is obvious, but there are many less obvious embedded systems. Your car contains dozens of processors, and even more in its audio system. For instance, each airbag is controlled by its own computer chip. Digital cameras, DVD players, medical equipment, and even home appliances contain dedicated processors.

The overwhelming majority of processor chips are used in embedded systems. Less than 1% of processor chips are used in the familiar desktop computer!

QUESTION 10:

Your automobile engine is controlled by a computer. Do you think that this computer has:

- Input devices?
- Output devices?
- Main storage?
- Secondary storage?

Answer:

Does the computer controlling an automobile engine have:

- Input devices? **Yes** — various *sensors* throughout the car deliver data to the computer.
- Output devices? **Yes** — various *actuators* (that perform actions) are controlled by the computer.
- Main storage? **Yes** — ROM holds the program and there must be enough RAM to hold the current data and results.
- Secondary storage? **No** — Hard disks do not hold up well next to a car engine.

Software

Computer software consists of both *programs* and *data*. Programs consist of instructions for the processor. Data can be any information that a program needs: character data, numerical data, image data, audio data, and countless other types. The distinction between programs and data is not as clear-cut as you might think, however.

Fundamental Idea: Both programs and data are saved in computer memory in the same way. The electronics of computer memory (both main memory and secondary memory) make no distinction between programs and data.

The insight that both programs and data can be saved using the same electronic methods is an important concept in computer science. Computer systems use memory for either programs or data, as needed.

QUESTION 11:

Which of the following can be saved on a disk?

- A game program
- Accounting data (numbers)
- Audio data
- Graphics data
- Digital photographs
- Video data

Answer:

Everything on the list can be stored on a disk (or on any other kind of computer memory, for that matter).

Types of Programs

There are two categories of programs. **Application** programs (usually called just "applications") are programs that people use to get their work done. Computers exist because people want to run these programs. **Systems** programs keep the hardware and software running together smoothly. The difference between "application program" and "system program" is fuzzy. Often it is more a matter of marketing than of logic.

The most important systems program is the **operating system**. The operating system is always present when a computer is running. It coordinates the operation of the other hardware and software components of the computer system. The operating system is responsible for starting up application programs, running them, and managing the resources that they need. When an application program is running, the operating system manages the details of the hardware for it. For example, when you type characters on the keyboard, the operating system determines which application program they are intended for and does the work of getting them there.

Some embedded systems do not use an operating system, but run their programs directly on the processor.

Modern operating systems for desktop computers come with a **user interface** that enables users to easily interact with application programs (and with the operating system itself) by using windows, buttons, menus, icons, the mouse, and the keyboard. Examples of operating systems are Unix, Linux, Android, Mac OS, and Windows.

Application Programs	Systems Programs
Word processors Game programs Spreadsheets Data base systems Graphics programs Web browsers	Operating system Networking system Programming language software Web site server Data backup

QUESTION 12:

(Thought question:) Imagine that you have just turned on your computer and have not yet started any application. Are any programs running?

Answer:

Yes—the operating system is running. Mostly it is managing the user interface, waiting for some input to tell it what to do.

Operating Systems

An operating system is a complex program that keeps the hardware and software components of a computer system coordinated and functioning. It is like the owner of a small shop, who keeps everything in order by attending to customers, accepting deliveries, stocking the shelves, doing the bookkeeping, and so on. The shopkeeper must promptly attend to tasks as they arise. Without the shopkeeper the shop could not function.

Most computer systems can potentially run any of several operating systems. For example, most Pentium-based computers can run either Linux or a Windows operating systems. Usually only one operating system is installed on a computer system, although some computers have several. In any case, only one operating system at a time can be in control of the computer system. The computer user makes a choice when the computer is turned on, and that operating system remains in control until the computer is turned off.

QUESTION 13:

(Review:) What component of a computer system holds the operating system when the computer is not running?

Answer:

The hard disk. The hard disk is used for long-term storage of all software, including the operating system.

Starting a Program

When a computer is first started, the hardware automatically loads the operating system and starts it running. This process is called **booting**. The reason for this odd term is that the operating system is itself involved in getting itself running—a process that is like someone "pulling themselves up by their bootstraps". Once the operating system is running, it is used to start up application programs.

Here is a (simplified) list of what happens when the user (you) starts up an application. Assume that the operating system (OS) is already running.

1. The user asks to run an application.
 - This is done by clicking on an icon, making a menu choice, or by other means.
2. The OS determines the name of the application.
3. The OS finds the files on the hard disk where the application and its data are stored.
4. The OS finds an unused section of main memory that is large enough for the application.
5. The OS makes a copy of the application and its data in that section of main memory.
 - The software on the hard disk is unchanged; main memory holds a *copy* of what is on disk.
6. The OS sets up resources for the application.
7. Finally, the OS starts the application running.

As the application runs, the OS is there in the background managing resources, doing input and output for the application, and keeping everything else running.

QUESTION 14:

(Thought Question:) When an application stops running, do you think that it is copied from main memory back to the hard disk?

Answer:

No. The hard disk already has a copy, and that copy will be used the next time the application is run.

Networks

A computer **network** consists of two or more computers connected so that they can exchange data and programs. When a computer is a member of a network, the programs it runs and the data it uses can be on the hard disk of some other computer on the network. In business and industrial settings, most computers are on a network. The operating system that runs on a networked computer must manage its share of the network (along with managing all its other responsibilities). The operating system is able to find programs and data that are stored on other network computers, and copy them into its own main memory.

In a **local-area network** only a few dozen computers are connected together, usually all located within the same building. Each computer has a **network address** that the other computers use to access it. Usually the computers share a printer. There may be an especially powerful computer called a **server** whose hard disk holds application programs and data that the other computers are expected to need.

Each computer in a network has a **network interface card**. This is an input/output device that sends and receives data over cables. The network interface cards of computers on a network are connected together with cables.

QUESTION 15:

Can a local area network be implemented using radio waves instead of cables?

Answer:

Yes.

Wide-Area Networks

Large organizations need to connect many more computers than can be handled with a local area network. A **wide-area network** can connect thousands of computers together over great distances. The long distance connections are made by using optical fiber, telephone lines, microwave radio, and satellite communications. Each computer in the network has a network address (as with local-area networks) to uniquely identify it.

Wide-area networks use a variety of special hardware to manage the flow of data. When two computers share data, this hardware makes it appear that the two computers are connected together directly. In reality, there may be dozens of network devices between the two computers.

All these devices use the same method for dealing with data. Without a common method of dealing with data, a large network would become a hopeless muddle. An agreement about how to represent and transmit data over a network is called a **protocol**. Usually large networks use a protocol called TCP/IP (for transmission control protocol / internet protocol).

QUESTION 16:

(Thought Question:) When two computers communicate must they both use the same protocol?

Answer:

Yes. There are several protocols available, but when two computers communicate they must agree on which one to use.

Internet

The **Internet** consists of many networks that have been connected together to form one huge worldwide network. Even on this huge network, each computer must have a unique network address, called an **IP address**, much like each telephone in the world has a unique telephone number (including the country code and area code).

Here a typical IP address:

149.152.21.203

IP addresses actually are 32-bit binary numbers. Networking equipment uses these addresses to route information over the network. The above example shows the standard way of writing these bits using decimal digits.

But even when written in decimal the address is not clear to humans. One of the features of the World Wide Web is that it allows humans to use **computer names** rather than numbers. Here is a typical computer name:

chortle.ccsu.edu

This computer name corresponds to the above IP address. When you use it in a Web browser it is converted into the 32 binary digits of the computer's IP address.

As an experiment, enter the above IP address into the address field of your browser and hit return. Now do the same with the computer name. You will get to the same web page with each.

QUESTION 17:

(Thought Question:) With so many computers connected to the Web, do you suppose that there is trouble giving each computer its own address?

Answer:

Yes. The Web is running out of addresses. The next generation of Internet hardware and software will fix this (and other) problems.

World Wide Web

Remember that important idea (discussed several pages back):

Fundamental Idea: Both programs and data are saved in computer memory in the same way. The electronics of computer memory (both main memory and secondary memory) make no distinction between programs and data.



Communications equipment makes no distinction between programs and data, either. It is all information as far as it is concerned, and all information is transmitted the same way. The Internet is like a worldwide package delivery service. It is concerned with moving packages from one address to another, without concern about what is in the packages.

The Internet provides the hardware and the information transmission protocols for the **World-Wide Web**. Data intended for the Web is transmitted over the Internet just like any data. What makes Web data special is that it is intended for Web browsers (such as the one you are probably looking at). A browser is a program that can read Web pages and display them in a nicely formatted way.

A Web page is a package of data that contains information on how it is to be displayed on a monitor. This information is given using a language called **Hypertext Markup Language (HTML)**. If you want to see the HTML that describes this page, look at the menu at the top of your Web browser, and select Tools/Web Developer/Page Source. For older browsers, select *View* in the menu then left-click on *Source*. This will bring up a new window with the HTML of this page in it. After you are done viewing, close the window by clicking on the close button in its upper right corner (the button marked with X).

QUESTION 18:

Can programs, as well as data, be exchanged over the Internet?

Answer:

Yes.

Recall the fundamental idea, that programs and data are saved in the same fashion. They both can be transmitted over the Internet in the same fashion, too.

Hyperlinks

One Web page is connected to another with a **hyperlink**. If you have been reading these notes over the Web, you have been linking between Web pages by using hyperlinks.

A Web browser usually displays a hyperlink in a distinguishing color (usually blue). When you click on it, the browser asks the operating system to get a particular Web page from another computer connected to the Internet.

The Web page to get is specified with a **uniform resource locator** URL. A URL specifies the exact computer (among all the Internet computers in the world) and the exact Web page on that computer.

To see some examples of URLs, keep watching the box at the top of your browser labeled "Address".

QUESTION 19:

When you click on the following hyperlink, will the URL in the address box change?

Answer:

Yes. The new Web page is specified with its unique URL.

End of the Chapter

This ends the general overview of computer systems. At this point, you should have an idea of what computer systems are all about. But you will likely be uncertain about quite a few things.

The following chapters will clarify some topics, and others will remain mystifying until later courses. For now, you may wish to review the following. Click on a subject that interests you to go to where it was discussed.

- Hardware and software.
- Major hardware components of a computer system
- Types of memory
- Characteristics of a hard disk.
- I/O devices
- Two categories of programs
- The operating system.
- How programs start running.
- Computer networks

The next chapter will discuss *analog* and *binary* signals and why binary signals are important in computer systems.

You have reached the end of the chapter.