

CSCI 120 Introduction to Computation

Ports and I/O devices (cont.) (draft)

Saad Mneimneh
Visiting Professor
Hunter College of CUNY

1 Input devices

An input device is any hardware component that allows users to enter data (programs, commands, and user responses) into a computer. Depending on the application and the particular requirements, the input device selected may vary.

1.1 Keyboard

This is the most basic input device. The keyboard contains keys users press to enter data into a computer. Desktop computer keyboards typically have from 101 to 105 keys. Keyboards for smaller computers, such as laptops, contain fewer keys.

All computer keyboards have a typing area that includes the letters of the alphabet, numbers, punctuation marks, and other basic keys. The keys originally were arranged to reduce the frequency of key jams on old mechanical typewriters. The first six letters on the top alphabetic line spell QWERTY, thus giving the name QWERTY keyboard. A DVORAK keyboard, by contrast, places frequently typed letters in the middle of the typing area. Despite the logical design of a DVORAK keyboard, all keyboards today are QWERTY.

Many keyboards also have a numeric keypad on the right side of the keyboard, two CTRL keys, two ALT keys, a set of arrow and additional keys between the typing area and the numerical keypad, and twelve function keys (enhanced keyboard). The command associated with a function key may vary, depending on the program with which we are interacting.

1.2 Pointing devices

A pointing device is an input device that allows a user to control a pointer on the screen. In a graphical user interface, such pointer is usually a small symbol whose location and shape change as a user moves the pointing device. A pointing device can be used to move the insertion point of typed text, select a portion of the text, select graphics, click buttons, icons, links on a web page, and menu commands.

1.2.1 Mouse

The mouse was invented by Douglas Engelbart in 1964 and was refined by Xerox and employed for the first time with Apple Macintosh computer in 1970s. It was mass produced in the mid 1980s. It is now the most widely used pointing

device. With a mouse, the user controls the movement of the pointer, often called the mouse pointer in this case. The mouse pointer moves on the screen mimicking the mouse movements. The mouse has buttons that can be clicked once the mouse pointer is at the desired spot, to perform various actions.

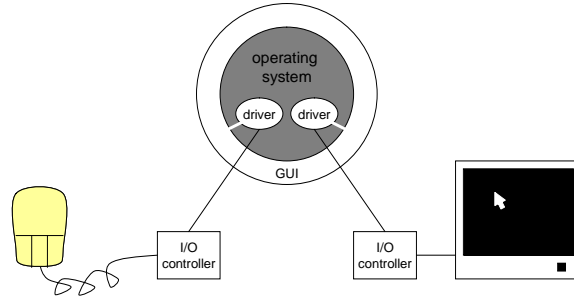


Figure 1: This shows how the mouse can control the pointer on the screen. The operating system (OS) is a software program that is continuously running to coordinate the activity of the whole machine, e.g. Windows; for instance, it's the OS that loads other programs in memory. The OS has a number of components among which are the Graphical user interface (GUI) and drivers for I/O controllers.

There are different types of mice:

mechanical mouse

A mechanical mouse has a rubber or metal ball on its underside. When the ball rolls in a certain direction, electronic circuits in the mouse translate that movement into signals the computer can interpret. A mechanical mouse is usually placed on a mouse pad to exert enough friction to ensure that the ball rolls as the mouse moves.

optical mouse

An optical mouse, by contrast, has no moving mechanical parts inside. Instead, an optical mouse uses devices that emit and sense light to detect the mouse movement. While some use optical sensors, others use laser. An optical mouse can be placed nearly on all types of surfaces, but reflecting surfaces such as mirrors will confuse the mouse! An optical mouse is more precise than a mechanical one and does not require cleaning as does a mechanical mouse.

wireless mouse

The mouse, whether mechanical or optical connects, to a computer with a cable that attaches to a serial port, mouse port, or USB port. But it can also be wireless, a battery powered mouse that transmits data using wireless technology (e.g. Bluetooth).

1.2.2 Trackball

Similar to a mechanical mouse, a trackball is a stationary pointing device with a ball on its top or side. The ball in most trackballs is about the size of a Ping-Pong ball. To move the pointer using a trackball, the ball is rotated with the thumb, the fingers, or the palm. The pointer moves on the screen mimicking the movements of the ball.

1.2.3 Touchpad

A touchpad is a small, flat, rectangular pointing device that is sensitive to pressure and motion. To move the pointer using a touchpad, the fingertip slides across the surface of the pad indicating the direction of the movement. Some touchpads have one or more buttons around the edge of the pad that work like mouse buttons. On most touchpads, we can also tap the pad's surface to imitate mouse clicking. Examples of touchpads include laptop touchpads and control pads on audio players such as iPod.

1.2.4 Pointing stick

A pointing stick is a pressure sensitive pointing device shaped like a pencil eraser that is positioned between keys on a keyboard (usually on IBM Thinkpad laptops). The pointer on the screen moves in the direction in which the pointing stick is pushed.

1.2.5 Light pen

A light pen is a handheld input device that can detect the presence of light. To select objects on the screen, the light pen is pressed against the surface of the screen or pointed at the screen, then a button on the pen is pressed (much like a mouse).

1.2.6 Touch screen

A touch screen is a touch-sensitive display device. Users can interact with these devices by touching areas on the screen, e.g. ATMs. Because touch screens require a lot of arm movements, we do not usually enter large amount of data using a touch screen. Instead, we touch words, pictures, numbers, letters, etc...

1.2.7 Pen input

Mobile users often enter data with a pen type device. With pen input, users write, draw, and tap on a flat surface to enter input. The surface may be a monitor, a screen, a special type of paper, or a graphics tablet. Two devices used for pen input are the stylus and the digital pen. The **stylus** is a small metal or plastic device that looks like a tiny ink pen but uses pressure instead of ink. A **digital pen**, which is slightly larger than a stylus, is available in two forms: some are pressure-sensitive, others have built in digital cameras. Examples where a pen input is used include Palm Pilots (stylus), PDAs, smart phones, and tablet PCs (digital pen).

1.3 Audio input

Audio input is the process of entering any sound into the computer such as speech, music, and sound effects. To enter high quality audio, the computer must have a sound card (the controller). Audio is entered via a microphone, CD/DVD player, and other devices, each of which plugs to a port on the sound card. Music can be entered using external MIDI devices such as electronic piano keyboards. In addition to being a port on the sound card, MIDI (Musical Instrument Digital Interface) is the electronic music industry standard that defines how music is represented electronically. Other types of audio input include voice. Special software programs, known as **voice recognition**, convert

the spoken words on the microphone to characters. Many programs, such as Microsoft Office, support voice recognition. For example, Microsoft Word can recognize dictated words and voice commands.

1.4 Digital camera

A digital camera allows users to take pictures and store the photographed images digitally. Generally speaking, a digital camera works in the following way:

- when shooting, light passes through the lens of the camera
- the image is captured using image sensors (sensitive to light)
- a charge coupled device (CCD) is used to generate an analog signal from the sensors
- the analog signal is converted to a digital signal by an analog to digital converter ADC
- a processor in the camera adjusts the quality of the image and then stores the digital image

The image is stored on a compact flash card or a memory stick (by SONY) or a smaller SD (Secure Digital) card. Later, the image can be transferred to the hard disk, by USB for instance, for editing if desired using various software applications.

A digital camera's resolution is measured in megapixels, i.e. million pixels (think of a pixel as a tiny dot, it is short for picture element). For example, a camera with an array of 2048×1536 sensors is commonly said to have 3.1 megapixels ($2048 \times 1536 = 3,145,728$). Since a pixel is just a logical unit of information (unless we specify its size), two important measures are to be considered when viewing or printing images, the the ppi (Pixels Per Inch) and the dpi (Dot Per Inch), respectively.

ppi is the more straightforward of the two measures. It describes just that: how many pixels an image contains per inch of distance in the horizontal and vertical directions. dpi may seem deceptively simple at first. The complication arises because a printer may require multiple dots in order to create a single pixel; therefore, a given number of dots per inch does not always lead to the same resolution. Using multiple dots to create each pixel is a process called *dithering*: A printer with a limited number of ink colors can play a trick on the eye by arranging dots into small patterns, thereby creating the perception of a different color if each dot is small enough. As a result, images almost always require more DPI than PPI in order to achieve the same level of detail.

The most widely used format for representing the image is JPEG, developed by the Joint Photographic Experts Group. Actually, JPEG only specifies how the image is translated into bytes, but does not specify how these bytes are stored. Another standard JFIF (JPEG File Interchange Format) specifies how to store these bytes in a file. Therefore, when we refer to a JPEG file, we implicitly mean a JPEG/JFIF. JPEG is the format that is mostly used for storing and transmitting photographs on the World Wide Web. Other famous image formats include BMP, GIF, and PNG. Here's a list that summarizes these format:

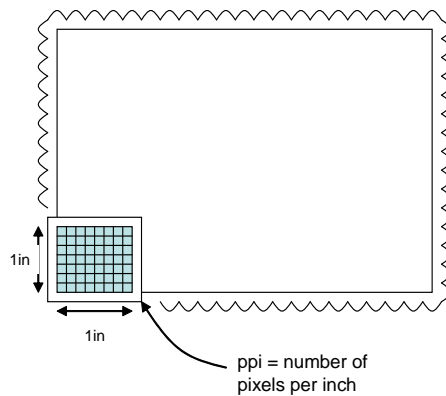


Figure 2: Megapixels and ppi, 300 ppi is good quality for viewing

- JPEG: this is what we mentioned above, and it is a commonly used standard method of **lossy compression** for photographic images. The image is compressed so that it occupies less space on the storage device. Lossy means that the original image cannot be retrieved from the compressed one (visual quality is degraded as a result of compression).
- BMP: Bitmap image. BMP images are usually uncompressed, so they are typically much larger than JPEG images. The word bitmap indicated that each pixel is represented by a number of bits, e.g. 24 bits. The bits determine the intensity of Red, Green, and Blue color. For instance, with 24 bits per pixel, the intensity of Red, Green, and Blue is represented by one byte each (8 bits for each), allowing values in the range 0 to 255. Therefore, a 640x480 pixel image will occupy 640x480x3 bytes, i.e. 900 KB almost 1 MB (compared to a few KB for a JPEG image). Bitmap formats are usually system dependent, e.g. Windows Bitmap is different that Unix Bitmap.
- GIF: The Graphical Interchange Format was developed by CompuServ in 1987. It is a bitmap image format, so each pixel is represented using a number of bits. However, it is compressed using the LZW (Lempel-Ziv-Welch) compression algorithm which was covered by U.S. Patent 4558302 owned by Unisys Corporation. This is a lossless compression. The file size is reduced without degrading the visual quality, provided that the image fits into 256 colors (note that with 24 bits per pixel, 16777216 colors are possible). The LZW patent expired on June 20, 2003. IBM also patented LZW but never enforced it. This patent will expire on August 11, 2006.
- PNG: Portable Network Graphics is a lossy compressed bitmap image. PNG was created to both improve and replace GIF. PNG is officially pronounced "ping". The motivation for creating the PNG format came in early 1995, after Unisys announced that it would be enforcing software patents on the LZW data compression algorithm used in the GIF format (the acronym PNG was originally recursive, standing for "PNG's Not GIF"). There were also other problems with the GIF format which made a replacement desirable, notably its limitation to 256 colors at a time when computers capable of displaying far more than 256 colors were becoming common. Although GIF allows for animation, it was decided that PNG should be a single-image format.

1.5 Other input devices

Other input devices include video cameras, scanners, and various reading devices (bar code reader). In particular, scanners are widely used to save documents electronically. The scanner produces an image of the document. Software applications known as OCR (Optical Character Recognition) can then convert the image into text (if document represents a text). OCR applications are 99% accurate.

2 Output devices

Output is data that has been processed into a useful form. Computers process data (input) into information (output). Users view output on a screen, print it, or hear it through speakers. We will concentrate on display devices and printers.

2.1 Display devices

A display device, or simply display, is an output device that visually conveys information such as text, graphics, and video. The most widely used displays include TVs and computer monitors. Some displays can be used as both. In the computer world, information on a display is often called *soft copy* due to the fact that it exists electronically and appears for a temporary period. With advances in technology, displays have become more and more flat and shallow. The term “flat screen” is often used these days. We will look at the different types of display devices: Cathode Ray Tube (CRT), Liquid Crystal Display (LCD), and Plasma.

2.1.1 CRT

A CRT display is one that contains a cathode ray tube (see Figure 3 below). A cathode ray tube is a large, sealed glass tube. The front of the tube is the screen (the shape of the tube explains why the traditional monitor has this shape).

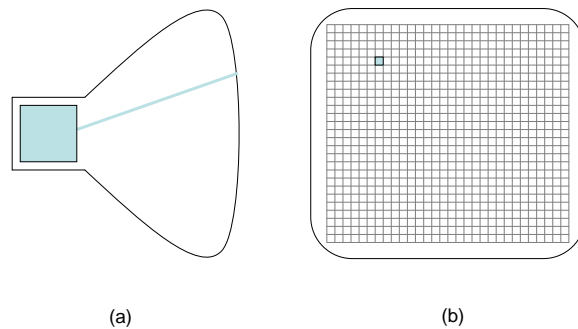


Figure 3: A CRT display. (a) Side view of the cathode ray tube. (b) The front view of the cathode ray tube, i.e. the screen.

Inside the CRT (at the narrow side in Figure 3(a), there is an electron gun. This electron gun fires electrons in the form of an electron beam that hits the back of the screen (hence the name cathode ray). The screen is made of tiny dots (usually phosphor, like the material covering the watch arms) that glow when hit by the electron beam (see Figure 3). The amount of illumination

is determined by the intensity of the beam. The electron gun continuously moves back and forth across the back of the screen hitting every dot with the appropriate intensity, thus producing the image.

But how do we see color? Each dot on the screen actually consists of three tiny dots of Red, Green, and Blue phosphor. When excited with heat, electron ray, or ultra violet light, phosphor emits a visible light depending on its color (try to place your watch under the light and then watch it glow in the dark).

Therefore, instead of one electron gun, three electron guns are used, one for each of the three colors. As a result, each dot displays the color obtained by a combination of different intensities of Red, Green, and Blue.

Resolution is the number of horizontal and vertical pixels in a display device. For example, a display device that has a resolution of 1600×1200 displays up to 1600 pixels horizontally and 1200 pixels vertically, for a total of 1920000 pixels. Most CRT monitors support a variety of screen resolutions. Standard CRT monitors today usually display up to a maximum of 1800×1440 pixels, with 1280×1024 being often the norm. Some CRT monitors can display 2048×1536 pixels or more.

Pixel pitch, sometimes called *dot pitch* is the distance in millimeters between pixels on display device. Of course, the lower the number, the sharper the image. Text created with a smaller dot pitch is easier to read; therefore, to minimize eye fatigue, a CRT with a dot pitch of 0.27 millimeters or lower is typically used.

The electron gun inside a CRT “draws” the image on the entire screen many times per second so that the phosphor dots do not fade. The number of times the image is drawn per second is called the *refresh rate* or the *scan rate*. To maintain a constant flicker-free image, the refresh rate must be fast enough. A high quality CRT monitor provides a vertical refresh rate of at least 68 Hz. This means the image on the screen redraws itself vertically 68 times per second.

Many CRT displays use an analog signal for the cathode ray tube. Therefore, the controller (known as graphics card) needs to have a DAC (Digital to Analog Converter) to convert the digital signal to an analog signal. Although CRTs today provide high quality images, due to the cathode ray tube, CRT displays are much bulkier than LCDs which can be very shallow.

2.1.2 LCD

As mentioned previously, LCD stands for Liquid Crystal Display. LCDs are found in laptop computers, digital clocks and watches, microwave ovens, CD players, MP3 players, calculators, and many other electronic devices. They are that common because they offer some real advantages over other display technologies. For instance, they are thinner and lighter and draw much less power than CRTs.

The name liquid crystal sounds like a contradiction. A liquid crystal is a substance that can exist in a state that is sort of like a liquid and sort of like a solid. But a liquid crystal is closer to a liquid state than a solid. It takes a fair amount of heat to transform a suitable substance from a solid into a liquid crystal, and only a little more heat to turn that same liquid crystal into a liquid.

A particular type of liquid crystals, called twisted nematic, are naturally twisted. Applying an electric current to these liquid crystals will untwist them to varying degrees. This forms the basis for LCDs. Before we understand how this twisting/untwisting feature makes the LCD work, we need to talk about light polarization.

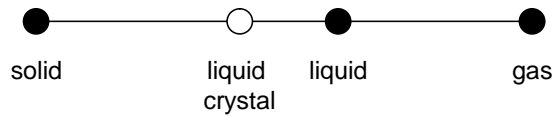


Figure 4: liquid crystal state

Light waves vibrate in all directions. When light hits an suitable surface (e.g. polarizing filter), only the component vibrating perpendicular to the direction of the surface molecules goes through. The rest is absorbed or reflected by the surface. The light is said to be polarized (see Figure 5 below).

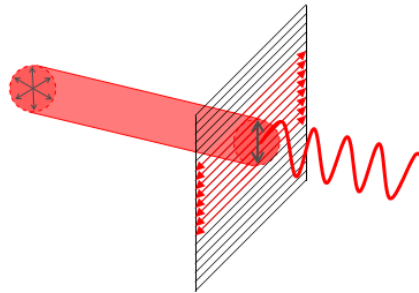


Figure 5: Polarization

The LCD works as follows. As light strikes a first filter, it is polarized. The filter is followed by twisted nematic. The molecules in each layer of twisted nematic then guide the light they receive to the next layer. As the light passes through the liquid crystal layers, the molecules also change the light's plane of vibration to match their own angle. When the light reaches the far side of the liquid crystal substance, it vibrates at a 90 degree angle with respect to its initial polarization. Therefore, adding a second filter will block the light (see Figure 6 below).

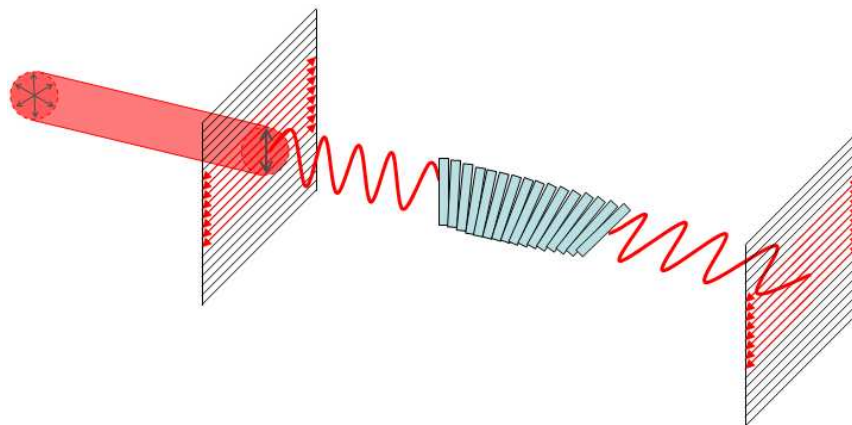


Figure 6: Light is blocked by twisted nematic

When an electric current is applied, the twisted nematic is gradually untwisted until all molecules are aligned, allowing the light to maintain the same

initial polarization, hence passing through the second filter too.

Color is obtained by using three light beams (Red, Green, and Blue) with different intensities for every pixel on the screen.

Two basic technologies of LCDs are used. Passive matrix or Active matrix (also known as Thin Film Transistor TFT). In passive matrix, when voltage is applied to untwist one pixel, the pixels around it also partially untwist. This makes images appear fuzzy and lacking in contrast. In active matrix, a separate transistor is used to apply voltage to each liquid crystal pixel, resulting in faster response, better accuracy, and higher quality display. Generally speaking, LCDs are sensitive to temperature and have a slow response (observe your cell phone in a cold weather). A newer type of TFT, known as organic LED (OLED), uses organic molecules that produce even brighter and easier to read displays than traditional TFT displays.

A 17 inch LCD typically has a resolution of 1280×1024 . A 20 inch LCD typically has a resolution of 1600×1200 .

The response time of an LCD monitor is the time in milliseconds that it takes to turn a pixel on or off (twisting/untwisting). The response time ranges from 8 to 25 milliseconds.

Brightness of an LCD monitor is measured in **nits**. A nit is a unit of visible light intensity equal to one candela (formerly called candlepower) per square meter. The candela is the standard unit of luminous intensity. LCD displays and other types of displays today range from 200 to 350 nits. The higher the nit, the brighter the image.

The contrast ratio describes the difference in light intensity between the brightest white and the darkest black that can be displayed on an LCD. Contrast ratios today range from 400:1 to 800:1. Higher contrast ratios represent colors better.

Unlike most CRT displays, LCD displays use digital signals to produce the image. Therefore, to display the highest quality images, an LCD must connect to the DVI (Digital Video Interface) port on the graphic card, which enables digital signals to transmit directly to the LCD. Current video standards include SVGA (Super Video Graphics Array), which has a resolution of 800×600 ; XGA (eXtended Graphics Array), a resolution of 1024×768 ; SXGA (Super XGA), a resolution of 1280×1024 ; and UGA (Ultra XGA), a resolution of 1600×1200 .

Name	Resolution
CGA	320x200
EGA	640x350
VGA	640x480
SVGA	800x600
XGA	1024x768
SXGA	1280x1024
UXGA	1600x1200

2.1.3 Plasma display

The plasma technology makes it possible to build displays as thin as 6 inches. Similar to CRT, the basic principle of plasma displays is to illuminate tiny colored phosphor dots with different intensities to form the image. Unlike CRT, however, plasma technology uses gas to create ultra-violet light that excites the phosphor pixels.

Originally, the gas atoms are neutral: the number of electrons balances the number of protons. By establishing an electrical voltage across the gas,

free electrons are introduced. These free electrons collide with the gas atoms and knock loose other electrons. With a missing electrons, an atom becomes positively charged, i.e. an ion. With presence of ions and electrons, the gas becomes plasma.

Having opposite charges, ions and electrons in a plasma are constantly bumping each other. These collisions excite the gas atoms causing them to release photons of energy, usually in the form of an ultra-violet light (non-visible). This release of energy is due to an electron falling from a high energy level to a low energy level (see Figure 7 below).

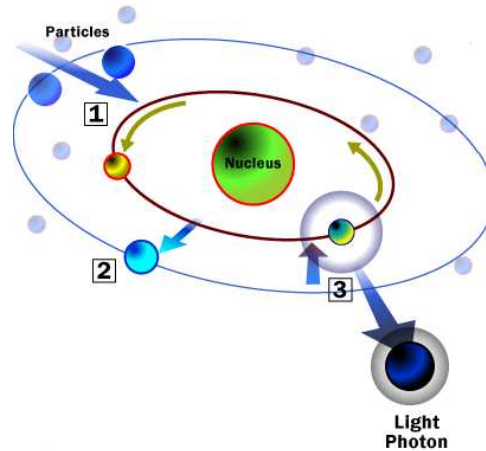


Figure 7: Plasma concept: (1) Collision with moving particle. (2) Electron jumps to a higher energy level. (3) Electron falls back to its original energy level, releasing extra energy.

The ultra-violet light is picked up by a colored phosphor (Red, Green, or Blue) which emits visible light when subjected to the ultra-violet light. As before, by varying the electrical voltage across the gas, different light intensities can be achieved.

2.2 Printers

A printer is an output device that produces text and graphics on a physical medium such as paper or transparency film. Printer information, called *hard copy*, exists physically and is a more permanent form of output than presented on a display device (*soft copy*). We will briefly look at three different printing technologies: dot matrix, ink jet, and laser.

2.2.1 Dot matrix (impact printing)

An impact printer forms characters and graphics on a piece of paper by striking a mechanism against an inked ribbon that physically contacts the paper. Impact printers characteristically are noisy because of this striking activity. These printers commonly produce *near letter quality* (NLQ) output, which is print quality slightly less clear than what is acceptable for business letters. These printers may be used today for routine jobs such as printing mailing labels, envelopes, and invoices.

A common type of impact printer is the dot matrix printer. The dot matrix printer produces printed images when tiny wire pins on a print head mechanism strike an inked ribbon. When the ribbon presses against the paper, it creates dots that form characters and graphics.

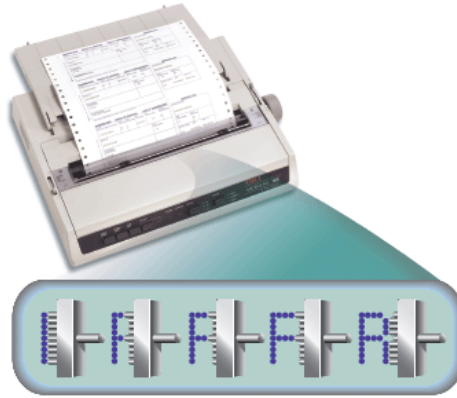


Figure 8: Dot matrix printer

Most dot matrix printers use continuous-form paper, in which thousands of sheets of paper are connected together end to end. The pages have holes along the sides to help feed the paper through the printer. The print head mechanism on a dot matrix printer contains 9 to 24 pins, depending on the manufacturer and the printer model. A higher number of pins means the printer prints more dots per character, which results in a higher print quality. The speed of a dot matrix printer is measured by the number of characters printed per second (cps). Most dot matrix printers have a speed ranging from 300 to 1100 cps.

2.2.2 Ink jet

An ink jet printer is a type of non impact printer that forms characters and graphics by spraying tiny drops of liquid ink onto a piece of paper. The print head mechanism in an ink jet printer contains ink filled print cartridges. Each cartridge has 50 to several hundreds small ink holes or nozzles. Here's how an ink jet works in principle:

1. A small resistor heats the ink, causing the ink to boil and form a vapor bubble.
2. The vapor bubble forces the ink through the nozzle.
3. The ink then drops onto the paper.

Ink jet printers have become a popular type of color printer for use at home. These printers normally use individual sheets of paper stored in one or two removable or stationary trays. As with many other input and output devices, one factor that determines the quality of an ink jet printer is its resolution. Printer resolution is measured by the number of *dots per inch* dpi. With an ink jet printer, a dot is just a drop of ink (better quality). A higher dpi, therefore, means that the drop of ink is smaller. Most ink jet printers can print from 600 to 4800 dpi. The speed of an ink jet printer is measured by the number of pages per minute (ppm) it can print. Most ink jet printers print from 3 to 26

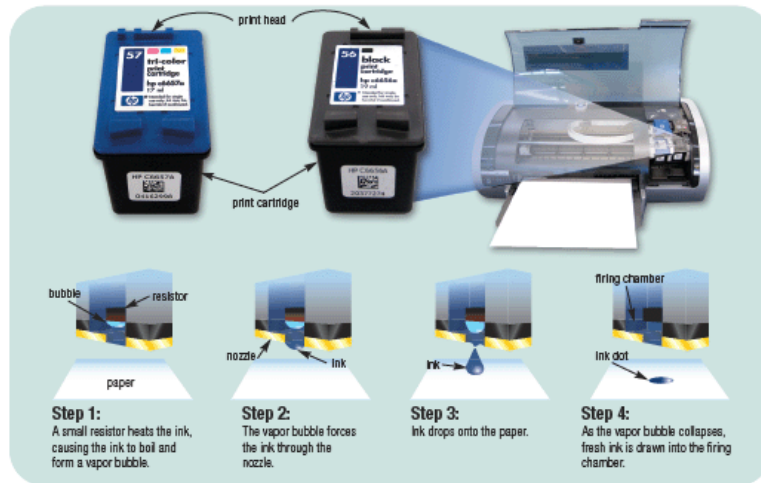


Figure 9: Ink jet printer

ppm, with graphics and colors printing at a slower rate. For instance, an ink jet printer may print 20 ppm of black text and only 15 ppm of color and/or graphics.

2.2.3 Laser

A laser printer is a high speed high quality non impact printer. Laser printers are available for black and white or color printing. A laser printer usually individual letter size sheets of paper stored in one or more removable trays. Some laser printers have built in trays that accommodate different sizes of paper. Laser printers print text and graphics in high quality resolutions, usually 1200 dpi for black and white printers and up to 2400 dpi for color printers. A laser printer for home and small office use typically prints at speeds of 15 to 50 ppm. Color laser printers print 4 to 27 ppm. Laser printers for larger business users print more than 150 ppm.

When printing a document, laser printers process and store the entire page before they actually print it. For this reason, laser printers are also called page printers. Storing a page before printing requires that the laser printer has a certain amount of memory in it, typically 544 MB of memory and a 20 GB hard disk. The printer is equipped with a laser beam (hence the name) and powdered ink, called *toner*. Here is how a laser printer prints the page:

1. After the user sends an instruction to print a document, a drum rotates as gears and rollers feed the sheet of paper.
2. The laser beam produces an image on the drum inside the printer by using a rotating mirror that deflects the laser beam across the surface of the drum.
3. The light of the laser alters the electrical charge on the drum wherever it hits, causing the toner to stick to the drum.
4. As the drum continues to rotate and press against the paper, the toner transfers from the drum to the paper.

5. A set of rollers uses heat and pressure to fuse the toner permanently on the paper as it gets out.

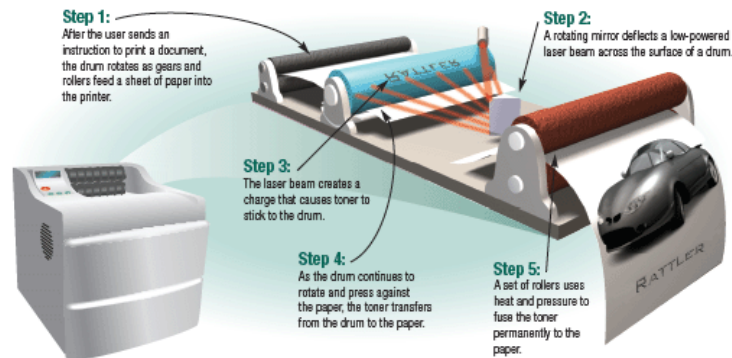


Figure 10: Laser printer

Laser printers use software that enables them to interpret a *page description language* (PDL), which tells the printer how to lay out the contents of a printed page. A common page description language is *PostScript* developed by Hewlett-Packard, a leading printer manufacturer. PostScript is a stack language. So what is a stack? A stack is a pile of objects. Imagine for instance that you are folding your shirts after they came out of laundry. When you put your shirts on top of each others, you form a stack. The last shirt you folded is the one that you can see. That shirt is the top of the stack and is the only shirt you can easily access. In general, a stack is a structure that supports two operations:

- push: add an object to the top of the stack
- pop: remove the top of the stack

For this reason, a stack is often called *Last In First Out* LIFO. A stack language works in almost the same way. Data is pushed onto the stack. When an instructions comes in, the operands for that instruction are popped (therefore, the instruction operates on the latest operands). The result of the instruction is pushed onto the stack. Another example of a stack language is the postfix calculator. The following figure illustrates an example use of postfix calculator.

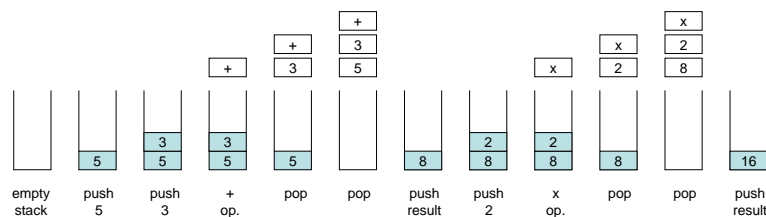


Figure 11: $5\ 3\ +\ 2\ *$ as computed by a postfix calculator

3 Storage devices

Storage devices, such as disk drives, serve as both input and output devices. We discussed the various types of storage devices in Lecture 4.