25 Semiconductor Devices and Information Technology

KEY IDEA

Semiconductor devices enable us to transmit and process information quickly, reliably, and conveniently.

PHYSICS AROUND US . . . Keeping in Touch

Your best friend from high school attends college three states away, but you like to stay in touch. You turn on your computer, enter your e-mail system, and send him a message telling him about your week. Later in the day he sends you an e-mail back. It feels good to stay in touch when you're far away from friends.

Later that day you log on to the Web to do research for a term paper about government. Your search takes you to Web pages posted by organizations in many countries and on several continents. All this information, including photos, diagrams, and references to other sources, appears on your computer screen in seconds.

You don't give it a second thought, but all of that information is brought right into your room by a huge network of computers and electronic storage systems that spans the entire globe—a network made possible by the technology of semiconductors.

MICROCHIPS AND THE INFORMATION REVOLUTION

Every material has dozens of different physical properties. We've already seen how strength, electrical conductivity, and magnetism all result from the properties of individual atoms and how those atoms bond together. We could continue in this vein for many more chapters, examining optical properties, elastic properties, thermal properties, and so on. But such a treatment would miss another key idea about materials: understanding how atomic interactions affect the properties of materials can lead to the development of new materials, and new materials can lead to new technologies that change society.

Of all the countless new materials discovered in the twentieth century, none has transformed our lives more than silicon-based semiconductors. In personal computers, auto ignitions, portable radios, sophisticated military weaponry, and countless other devices, microelectronics is a hallmark of our age. Indeed, semiconductors have fundamentally changed the way that we communicate one of society's most precious resources—information. The key to this revolution is our ability to fashion complex crystals atom by atom from silicon, a material that is produced from ordinary beach sand.

Doped Semiconductors

The element silicon by itself is not a very useful substance in electric circuits. What makes silicon useful and has driven our modern microelectronic technology is a process known as doping. **Doping** is the addition of a minor impurity to an element or compound. The idea behind silicon doping is simple. When silicon is melted before being made into circuit elements, a small amount of some other material is added to it. One common additive is phosphorus, an element that has five valence (bonding) electrons, as opposed to the four valence electrons of silicon.

When the silicon crystallizes to form the structure shown in Figure 25-1*a,* the phosphorus is taken into the crystalline structure. However, of the five valence electrons in each phosphorus atom, only four are needed to make bonds to

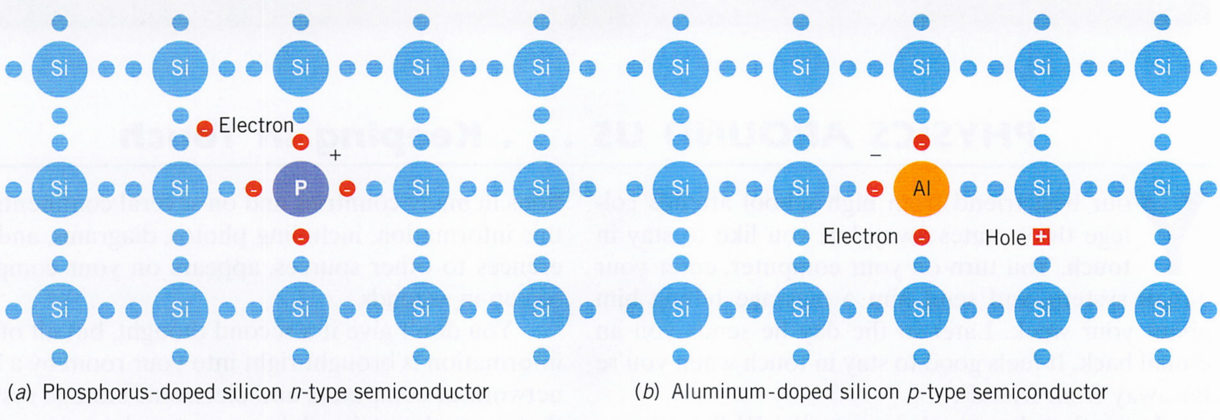


FIGURE 25-1. (*a*) Phosphorus-doped silicon *n*-type semiconductors and (*b*) aluminum-doped silicon *p*-type semiconductors are formed from silicon crystals with a few impurity atoms.

silicon atoms in the crystal. The fifth electron is not locked in at all. In this situation, it does not take long for the extra electron to be shaken loose and to wander off into the body of the crystal. This action has two important consequences: (1) conduction electrons are introduced into the material, and (2) the phosphorus ion that has been left behind has a positive charge. A semiconductor doped with phosphorus is said to be an "*n*-type semiconductor," because the moving charge is a negative electron.

Alternatively, silicon can be doped with an element such as aluminum, which has only three valence electrons (Figure 25-1*b*). In this case, when the aluminum is doped into the crystal structure there is one less valence electron compared to the silicon atoms it has replaced in the crystal. This missing electron—a hole—creates a material that can now more easily carry an electric current. The hole need not stay with the aluminum; it is free to move around within the semiconductor, as described in Chapter 24. Once it does so, the aluminum atom, which has now acquired an extra electron, has a negative charge. This type of material is called a "*p*-type semiconductor," because a positive hole—a missing negative electron—acts as the moving charge.

Diodes

You can understand the basic workings of a microchip by conducting an experiment in your mind. Imagine taking a piece of *n*-type semiconductor and placing it against a piece of *p*-type semiconductor. As soon as the two types of material are in contact, how will electrons move?

Near the contact, negatively charged electrons will diffuse from the *n*-type semiconductor over into the *p*-type, while positively charged holes will diffuse back the other way. Thus, on one side of the boundary will be a region where negative aluminum ions—ions locked into the crystal structure by the doping process—acquire an extra electron. Conversely, on the other side of the boundary is an array of positive phosphorus ions, each of which has lost an electron but is nonetheless locked into the crystal.

A semiconducting device such as this—formed from one *p* and one *n* region—is called a **diode** (see Figure 25-2). Once a diode is constructed, a permanent electric field tends to push electrons across the boundary in only one direction, from the *n*-type side to the *p*-type side. As electrons are pushed "with the grain" in the diode, from negative to positive, the current flows through normally. When the current is reversed, however, the electrons are blocked from going through by the presence of the built-in electric field. Thus the diode acts as a one-way gate, allowing the electric current through in only one direction.

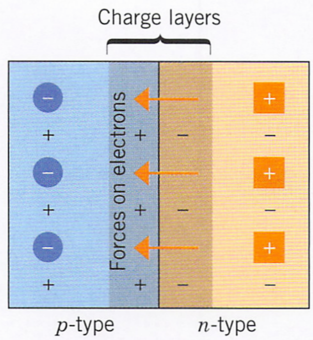


FIGURE 25-2. A semiconductor diode consists of an *n*-type region and a *p*-type region. Electrons in this diode can flow easily from the negative to the positive region. The built-in electric field, labeled E, blocks electrons from flowing the opposite way. The result is a one-way valve for electrons.

Semiconductor diodes have many uses in technology. One use can be found in almost any electronic device that is plugged into a wall outlet. As we have seen in Chapter 18, electricity is sent to homes in the form of alternating current (AC). It turns out, however, that most home electronic devices such as televisions and stereos require direct current (DC). A semiconductor diode can be used to convert the alternating current into direct current by blocking off half of it. In fact, if you examine the inside of almost any electronic gear, the power cord leads directly to a diode and other components that convert pulsing AC into steady DC. A semiconductor diode used in this way is called a *rectifier* (see Chapter 18).

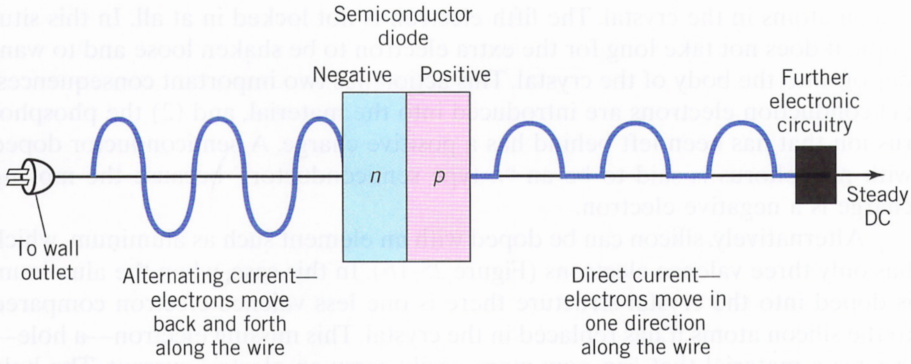


FIGURE 25-3. A diode converts alternating current from the wall into direct current in most electronic devices. Half of the alternating current passes through the diode, but the other half is blocked. Pulses of current from the diode enter a condenser, which stores the electrons and emits a steady DC current.

Figure 25-3 illustrates how a rectifier works. On the left, normal AC power enters into the system; recall that in AC, electrons flow first one way and then the other. If the current is in the half of the cycle with electrons flowing to the right in Figure 25-3, then these electrons pass through the rectifier. During the other half of the AC cycle, with electrons flowing to the left in the figure, electrons cannot move through the rectifier. Consequently, the output of the rectifier is a series of peaks, as shown, with the current in each peak moving in the same direction to the right. Further electronic circuitry then converts these peaks into a smooth DC voltage, as shown on the right. (Note that in most modern rectifiers, the current flowing in the wrong direction is not simply thrown away but converted to current flowing in the correct direction by more complex electric circuits.)

Connection

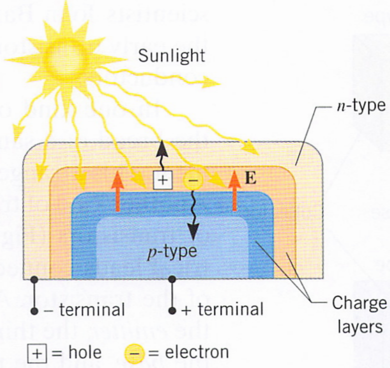
Photovoltaic Cells and Solar Energy

Recall from Chapter 22 that when some photoelectric materials absorb photons, electrons may be liberated from that material. This photoelectric effect is displayed by some semiconducting diodes, which may play an important role in the energy future of the United States. In such a device, called a *photovoltaic cell* (or solar cell), a thin layer of *p*-type material overlays a thicker layer of *n*-type. How might such a device operate to produce electricity?

Sunlight striking the top *n*-type layer shakes electrons loose from the crystal structure by the photoelectric effect. The electrons are attracted to the locked-in positive charges in the *n*-type material and repelled by the locked-in negative charges in the *p*-type material. They are then accelerated through the *n*–*p* boundary and pushed out into an external circuit. Thus, while the sun is shining, the photovoltaic cell acts in the same way as a battery. It provides a constant push for electrons and moves them through an external circuit. If large numbers of photovoltaic cells are put together, they can generate enormous amounts of current.

Photovoltaic cells have many uses today. Your hand calculator, for example, may very well contain a photovoltaic cell that recharges the batteries (it's the small dark band just above the buttons). Photovoltaic cells are also used in

(a)



(b)

(a) The Sun's energy is converted to electricity by photovoltaic panels at a southern California generating plant. (b) Photoelectric materials use the energy of photons to liberate electrons, which form an electric current.

regions where it is hard to bring in traditional electricity—for example, to pump water in remote sites or to provide electricity in backcountry areas of national parks. At the moment, it costs several times as much to get a watt of electricity from a solar cell as from the most expensive conventional generating plants, but this cost is falling as the technology improves and markets increase.

To understand how solar electricity might become a reality, you have to know a little about the way electricity is used in the United States. Certain demands exist all the time—people need to run computers, keep their streets and homes lit, run their subways, and perform countless other tasks at all hours of the day and night. Thus, utility companies need to supply a certain amount of electricity round the clock, day in and day out. This demand is called *base-load* electricity. Because the base-load generating plants operate for a large fraction of the time, base-load electricity is relatively cheap.

However, on certain days the demand for electricity soars—think of a hot August afternoon when everyone is running air conditioning, for example. Utility companies need to be able to meet this *peak-load* demand, but to do so they need to build plants that are not used all the time. Thus, peak-load electricity is more expensive to generate than base-load electricity.

Electricity demand peaks on hot summer afternoons, so experts who study energy policy have suggested that this is how solar energy will enter the electricity market. In this case, solar power would be competing with the more expensive peak-load generators, which makes the economic requirements somewhat less stringent. ●

The Transistor

The device that drives the entire information age and perhaps more than any other has been responsible for the transformation of our modern society is the **transistor.** Invented just two days before Christmas 1947 by Bell Laboratory

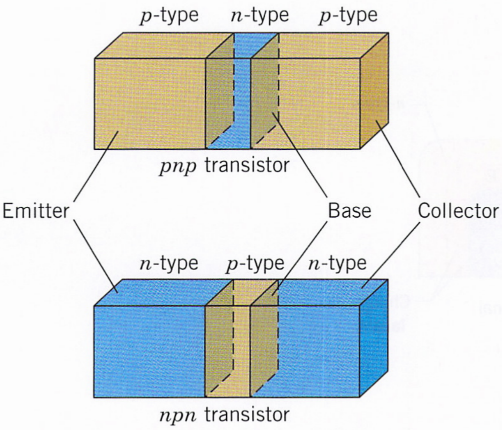


FIGURE 25-4. A *pnp* and an *npn* transistor.

scientists John Bardeen, Walter Brattain, and William Shockley, the early transistor was simply a sandwich of *n*- and *p*-type semiconductors.

In one kind of transistor, two *p*-type semiconductors form the bread of a sandwich, while the *n*-type semiconductor is the meat. This arrangement is abbreviated as the *pnp* configuration. Another kind of transistor uses the *npn* configuration. Both kinds of transistors (Figure 25-4) control the flow of electrons. Electrical leads connect to each of the three semiconductor regions of the transistor. An electric current goes into the region called the *emitter,* the thin slice of semiconductor in the middle is called the *base,* and the third semiconductor section is the *collector.*

Thus, the transistor has two built-in electric fields, one at each *p–n* junction. The idea of the transistor is that a small amount of electric charge running into or out of the base can change these electric fields—in effect, opening and closing the gates of the transistor. The best way to think of the transistor is to use a pipe that carries water as an analogy. The electric current that flows from emitter to collector is like water that flows through the pipe, and the base is like a valve in the pipe. A small amount of energy applied to turning the valve can have an enormous effect on the flow of water. In just the same way, a small amount of charge run onto the base can have an enormous effect on the current that runs through the transistor. These properties of a transistor lead to two of its most important uses—as an amplifier and as a switch.

The Transistor as Amplifier One use for the transistor is to amplify weak electric currents. For example, in your tape deck small electric currents are created when the magnetized tape is run past the tape heads. These small currents can be fed into the base of a transistor (Figure 25-5). As the number of electrons in the base increases, the main current flowing through the transistor decreases—in effect, the valve is closed a little. Similarly, when electrons are removed from the base, the current through the transistor increases—the valve is open. Thus,

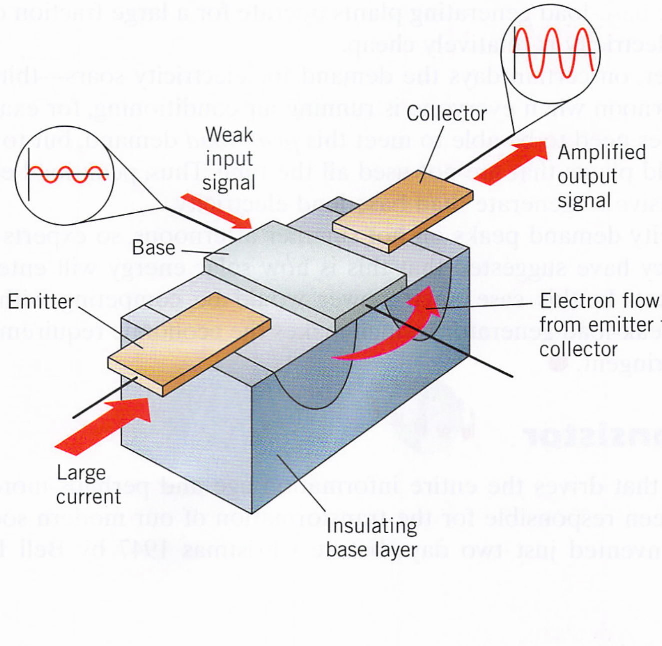


FIGURE 25-5. A transistor acting as an amplifier. A weak input signal sent to the base region modifies the stronger current passing from emitter to collector. The result is that the strong current carries an amplified version of the input signal.

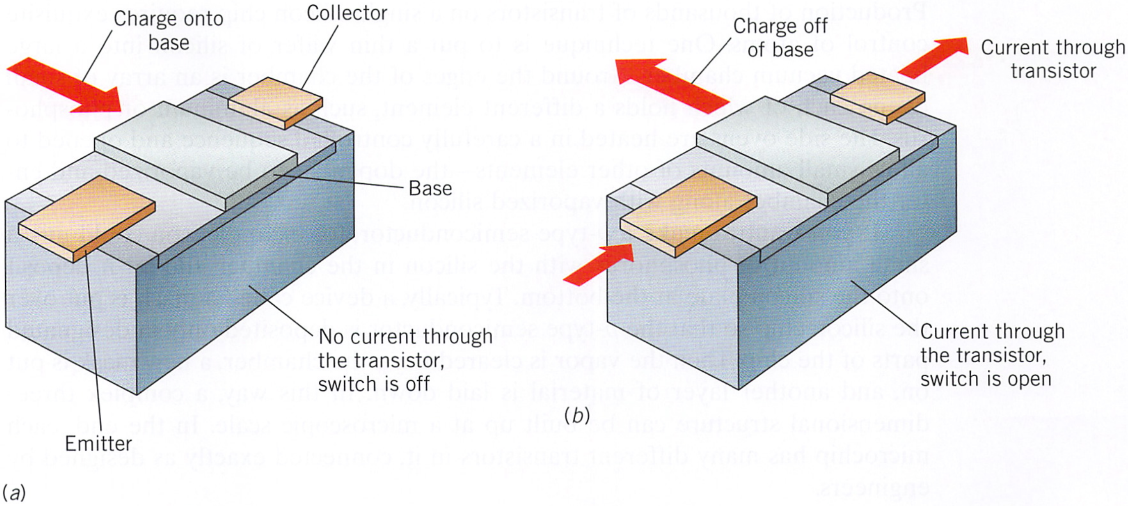


FIGURE 25-6. A transistor acting as a switch. (*a*) When current is supplied to the base, no current flows through the transistor; the switch is off. (*b*) When current does not flow to the base, a current can flow through the transistor; the switch is on.

the small signal from the tape can be impressed on the much larger current that is flowing from the emitter to the collector. It is this larger current that runs the speakers that produce the music you hear. A device that takes a small current and converts it into a large one is called an "amplifier."

The Transistor as Switch As important as the transistor's amplifying properties are, probably its most important use has been as a switch. If you run enough negative charge onto the base it can repel any electrons that are trying to get through—in effect, you can close the valve. Thus moving an electric charge onto the base shuts off the flow of current through the transistor, whereas running an electric charge off the base turns the current back on (Figure 25-6). In this manner the transistor acts as an electron switch and it can be used to process information in computers—surely the most important electronic device developed in the twentieth century.

Microchips

Individual diodes and transistors still play a vital role in modern electronics, but these devices have been largely replaced by much more complex arrays of *p*- and *n*-type semiconductors, called **microchips** or integrated circuits (Figure 25-7). Microchips may incorporate hundreds or thousands of transistors in one integrated circuit specially designed to perform a specific function. For example, an integrated circuit microchip lies at the heart of your pocket calculator or microwave oven control. Similarly, arrays of integrated circuits store and manipulate data in your personal computer, and they regulate the ignition in all modern automobiles.

The first transistors were bulky things, about the size of a golf ball. However, today a single microchip the size of a grain of rice can integrate hundreds of thousands of these devices. California's Silicon Valley has become a well-known center for the design and manufacture of these tiny integrated circuits.

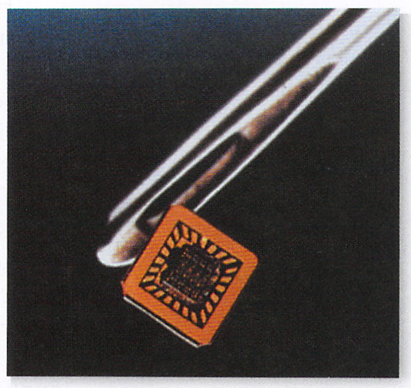


FIGURE 25-7. A microchip incorporates many complex circuits built into a single piece of silicon. For comparison, the eye of an ordinary sewing needle is shown in the background.

Production of thousands of transistors on a single silicon chip requires exquisite control of atoms. One technique is to put a thin wafer of silicon into a large heated vacuum chamber. Around the edges of the chamber is an array of small ovens, each of which holds a different element, such as aluminum or phosphorus. The side ovens are heated in a carefully controlled sequence and opened to allow small amounts of other elements—the dopants—to be vaporized and enter the chamber along with vaporized silicon.

If you want to make a *p*-type semiconductor, for example, you could mix a small amount of phosphorus with the silicon in the chamber and let it deposit onto the silicon plate at the bottom. Typically, a device called a mask is put over the silicon chip so that the *p*-type semiconductor is deposited only in designated parts of the chip. Then the vapor is cleared from the chamber, a new mask is put on, and another layer of material is laid down. In this way, a complex three-dimensional structure can be built up at a microscopic scale. In the end, each microchip has many different transistors in it, connected exactly as designed by engineers.

The reason that electronic devices have gotten so small is that engineers have gotten very good at creating smaller and smaller transistors on their microchips. This ability, in turn, depends on the ability to create finer and finer lines on the masks used in the fabrication process. Given the present rate of miniaturization, scientists have estimated that by the year 2040 the size of transistors will have reached a fundamental limit in which the lines will be one single atom across!

Physics in the Making

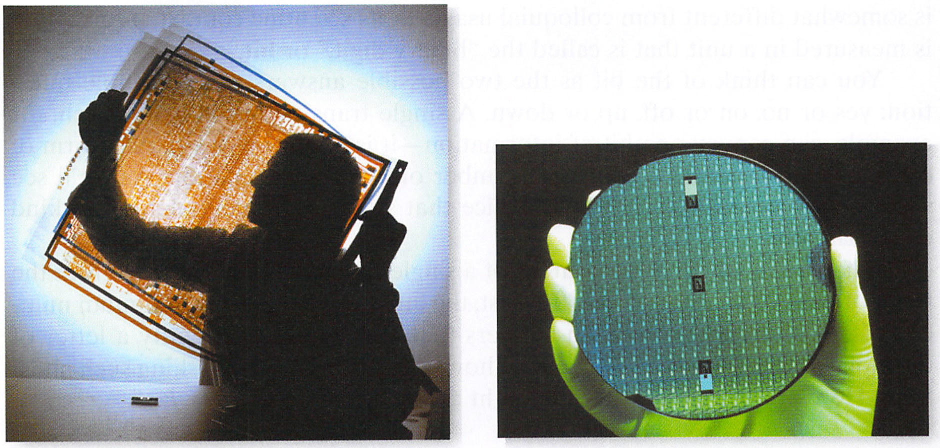
From the Transistor to the Integrated Circuit

When the transistor was invented in 1947, engineers quickly saw its advantages over the previous generation of electronic components, which were mostly vacuum tubes. Vacuum tubes in large numbers generated a lot of heat, requiring bulky cooling fans, and they burned out just like common lightbulbs. Transistors were small, reliable, cool, and efficient.

However, transistors presented problems of their own. Connections between transistors were made by the labor-intensive job of soldering wires into place under a microscope. And although the transistors themselves were reliable, it was all too easy for a wire to come loose. With electronics designers calling for more transistors in more complicated circuits, devices were hardly more practical than before.

The breakthrough came in 1958, when Jack Kilby, an electrical engineer at Texas Instruments, realized that you could create other circuit elements such as resistors and capacitors on the same slice of semiconductor material as the transistor. All the wires would be connecting elements on the same piece of material. His original prototype was a thin piece of germanium less than one-half inch long, containing five separate components connected by tiny wires. His patent application claimed that his *integrated circuits* could reduce the space taken up by electronic circuits by a factor of 60.

However, people still had to connect all the little wires—no easy task—and they were so small that they could break easily. In 1959 the second leap forward was made by Robert Noyce of Fairchild Semiconductors. Noyce realized that you could build the connections themselves into the circuit, using thin layers of metal



(*a*) (*b*)

(*a*) An electronics engineer examines an enlarged map of a complex integrated circuit. (*b*) Individual microscopic *n*- and *p*-type semiconducting regions decorate a silicon wafer containing integrated circuits.

to connect pieces of semiconductor. Noyce and colleagues at Fairchild developed a way of depositing different types of material on the base piece of silicon, covering up, or masking, areas where the next layer was not wanted. This technique became the basis for all modern microchip manufacture. Noyce went on to found the Intel Company, which is one of the largest chip manufacturers in the world.

Today, computers use microprocessor chips containing over 1 million transistors. For example, Intel's i486 microprocessor measures 0.414 by 0.649 inches and contains 1,180,235 transistors.

Kilby shared the Nobel Prize in Physics in 2000, one of the few prizes awarded for applied physics rather than pure research. Most science historians believe that Noyce would have shared the award if he had still been alive. (Noyce died in 1990, and Nobel Prizes are not awarded posthumously.) Nevertheless, few developments in physics have changed the world as much as their work in creating the microchip revolution. ●

INFORMATION

The single most important use of semiconducting devices is in the storage and manipulation of information. In fact, the modern revolution in information technology—the development of arrays of interconnected computers, global telecommunications networks, vast data banks of personal statistics, digital recording, and the credit card—is a direct consequence of materials science.

While it may appear strange to say so, almost all the media we normally consider as conveying information—the printed or spoken word, pictures, or music, for example—can be analyzed in terms of their information content and manipulated by the microchips we've just discussed. The term "information," like many words, has a precise meaning when it is used in the sciences—a meaning that is somewhat different from colloquial usage. In its scientific context, information is measured in a unit that is called the "binary digit," or **bit.**

You can think of the bit as the two possible answers to any simple question: yes or no, on or off, up or down. A single transistor used as a switch, for example, can convey one bit of information—it is either on or off. Any form of communication contains a certain number of bits of information. As we'll see shortly, the computer is simply a device that stores and manipulates this kind of information.

What is the information content of a single letter of the alphabet? From the point of view of an information theorist, the answer is simply the minimum number of questions with yes-or-no answers that are needed to identify a letter of the alphabet unambiguously. Let's see how you might go about asking such questions. Here are five questions you might ask to specify the letter *E*.

1. Is it in the first half of the alphabet? (yes)
2. Is it one of the first six letters of the alphabet? (yes)
3. Is it one of the first three letters of the alphabet? (no)
4. Is it *D*? (no)
5. Is it *E*? (yes)

From this simple example, you can see that you need the answers to at most five questions—5 bits of information—to specify unambiguously a single letter of the alphabet. (You might be lucky and guess the answer in fewer questions, but five questions are always enough to pinpoint any one of the 26 letters.) We can say, then, that the information content of a single letter of the alphabet is 5 bits.

As a matter of fact, $2 \times 2 \times 2 \times 2 \times 2 = 32$ different objects that can be specified using 5 bits of information. Thirty-two items is not enough, however, to also handle all the numbers, capitalization, punctuation marks, and other symbols used in writing. Six bits of information can specify $2 \times 2 \times 2 \times 2 \times 2 \times 2 = 64$ different items, and you could argue that everything you need to specify on a printed page can be included in those 64. Thus the information content of a single ordinary printed symbol is 6 bits. If you were using switches to store the information on a normal printed page, you would need six of them lined up in a row to specify each symbol.

The average word is six letters long, so the information content of a typical word is 36 bits. The average printed page of a novel contains about 500 words, corresponding to an information content of almost 20,000 bits. In this scheme, a 200-page book thus contains about 4 million bits, or 4 megabits, of information.

Historically, switches in computers were lumped together in groups of eight. Such a group is capable of storing 8 bits of information, or 1 **byte.** In terms of this unit, a 200-page book contains 500,000 bytes, or one-half of a megabyte.

This way of thinking about information—as a string of 0s and 1s—is ideally suited to a machine whose main working part is a transistor, which can be either on or off. Information represented in this way is said to be in **binary** form. It is a special case of information in **digital** form, which refers to a system that can have only a finite number of states.

Connection

Is a Picture Really Worth a Thousand Words?

Pictures and sounds can be analyzed in terms of information content, just as words can. Your television screen, for example, works by splitting the picture into small units called "pixels." In North America, the picture is split up into 525 segments on the horizontal and vertical axes, giving a total of about 275,000 pixels (in rounded numbers) for one picture on the TV screen. Your eye integrates these dots into a smooth picture. Every color can be thought of as a combination of the three colors red, green, and blue, and it is usual to specify the intensity of each of these three colors by a number that requires 10 bits of information to be recorded. (In practice, this procedure means that the intensity of each color is specified on a scale of about 1 to 1000.) Thus each pixel requires 30 bits to define its color, and the total information content of a picture on a TV screen is

$$275{,}000 \text{ pixels} \times 30 \text{ bits} = {\sim}8 \text{ million bits}$$

Thus it requires about 8 megabits, or 1 megabyte, to specify a single frame on a TV picture. We should note that a TV picture typically changes 30 times a second, so the total flow of information on a TV screen may exceed 200 million bits per second.

It thus would appear that a picture is worth much more than a thousand words. In fact, if a word contains 36 bits of information, then the picture will be worth 8 million bits per picture divided by 36 bits per word, which equals about 220,000 words per picture.

The old saying, if anything, underestimates the truth! ●

Connection

Playing Your CD

Much more than writing and pictures can be expressed digitally. For example, every time you play your favorite CD you rely on a digital representation of sound.

As we have seen in Chapter 15, sound is a pressure wave in the air, and the pressure of the air at your eardrum varies rapidly. To convert a variation like this to digital form, we sample the wave at equal time intervals and represent the pressure during that interval as a single number. The procedure, in essence, changes the variation in pressure from a smooth form to a stepped form. If the time intervals are short enough and the measurement of pressure sufficiently accurate, the ear is not able to tell the difference between the two curves.

When the recording for your favorite CD was made, the pressure wave was sampled 44,100 times each second. To convert the pressure to a digital number, the range of possible pressures was split into about 64,000 intervals and the averaged pressure of the wave at a given time was assigned to one of these intervals. As shown in Problem 11 at the end of this chapter, a 16-bit number can represent any one of these 64,000 intervals. Thus, the information content of a second of play on your CD is

$$\begin{aligned}\text{Information} &= (16 \text{ bits per interval}) \times (44{,}100 \text{ intervals per second})\\ &= 705{,}600 \text{ bits per second}\end{aligned}$$

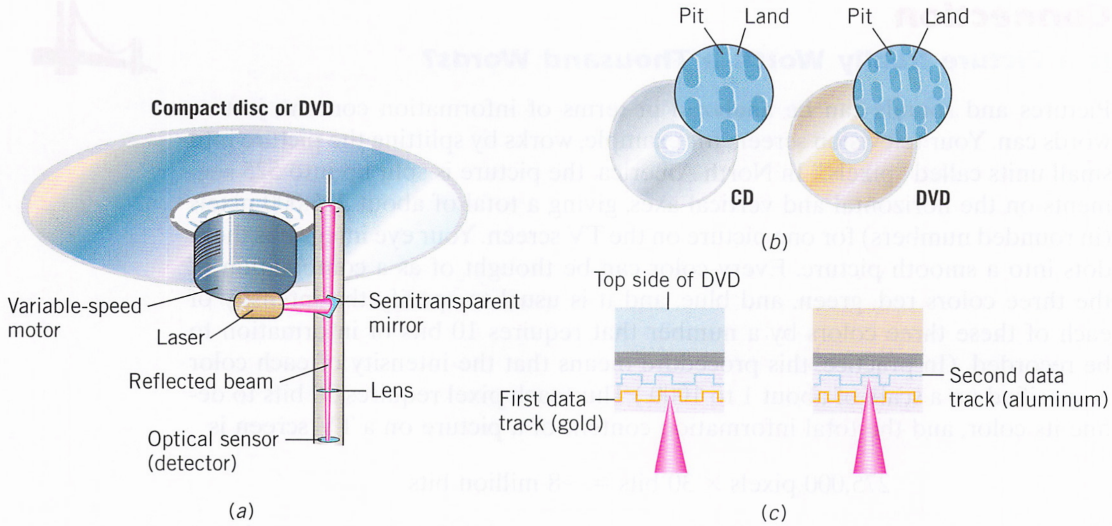


FIGURE 25-8. (*a*) A CD or DVD player bounces a laser beam off the bottom of the disc, where it is reflected to a detector or scattered by a pit. (*b*) A DVD has smaller pits and lands spaced closer together, so it can fit more data on the disc. (*c*) The DVD also has two layers of data tracks. The inner one is of aluminum; the second is of semitransparent gold. The combination gives DVDs their bronzelike color.

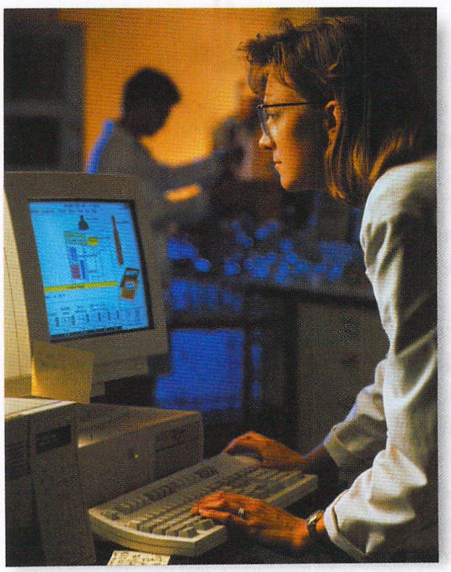
Thus, sound of CD quality contains more information than the written word, but less than a television picture.

To convert the digital signal back to sound, the disc that you put into your CD player has a series of pits in its surface (Figure 25-8). A laser beam plays on the disc's surface. If there is no pit on the surface, the beam is reflected into a detector. If there is a pit, the beam is deflected and goes unrecorded. Thus, the sequence of 0s and 1s of the digital signal are converted into a "beam-received" or "beam-not-received" signal by the laser. This signal, then, becomes the electric current that drives the speakers and produces the sounds you enjoy.

Over the past few years, digital video discs (DVDs) have become popular for storing and playing entire movies. A DVD is similar to a CD except that the pits are smaller and closer together. In addition, a DVD made with two layers of data tracks; the laser beam adjusts to a slightly higher power setting to read the inner track. These differences allow a DVD to store about ten times as much data as a CD. ●

Computers

A **computer** is a machine that stores and manipulates information. The information is stored in the computer in microchips, each of which incorporates many thousands of interconnected transistors that act as switches and carry information. In principle, a machine with a few million transistors could store the text for this entire book. In practice, however, computers do not normally work in this way. They have a *central processing unit (CPU)* in which transistors store and manipulate relatively small amounts of information at any one time. When the information is ready to be stored—for



In less than a quarter of a century, the computer has evolved from a specialized research aid to an essential tool for business and education.

example, when you have finished working on text in a word processor or writing a program to perform a calculation—it is removed from the CPU and stored elsewhere. For example, it might be stored in the form of magnetically oriented particles on a compact disc or a hard drive. In these cases, 1 bit of information is no longer a switch that is on or off but a bit of magnetic material that has been oriented either north pole up or north pole down.

The ability to store information in this way is extremely important in modern society. As just one example, think about making an airline reservation online. You enter a Web site with access to all commercial flight information and request a specific destination (at the lowest possible price). The site you entered polls the computers of several different airlines. Each of these computers stores strings of bits that represent different flights on different days, the seating assignments, ticket arrangements, and often the address and phone number of every passenger who will be flying on the particular day when you want to fly. When you change your reservation, make a new one, or perform some other manipulation, the information is taken out of storage, brought to the central processing unit, manipulated by changing the exact sequence of bits, and then put back into storage. This process—the storage and manipulation of vast amounts of data—forms the very fabric of our modern society.

You've undoubtedly noticed that the speed and information capacity of computers has increased astonishingly over the past few decades. In the mid-1980s the very best personal computers could store a few hundred thousand bits of information. By the mid-1990s typical personal computers held billions of bits. Today, lightweight laptop computers outperform the most advanced supercomputers of a decade ago with memories that exceed tens of billions of bits. These advances are primarily the result of many improvements in materials and their processing at the atomic scale—a field called *nanotechnology.* New fine-grained magnetic materials have greatly increased the capacity of information storage devices such as hard disks, while improved semiconductor processing techniques have dramatically reduced the size of individual *n*- and *p*-type domains. The result is smaller, more powerful computers. In this way, advances in materials science play a direct role in our lives.

Connection

Jim Trefil Gives His Car a Tune-Up

As a student, I acquired the first of a long string of Volkswagen Beetles. Now let me tell you, my friends, that was a sweet car! There were never any problems with the cooling system, for the simple reason that there wasn't any—the engine was cooled by the air flowing by. And almost any repair could be made by someone with reasonable mechanical ability and a set of tools. While in graduate school, I spent many happy hours under my car, adjusting this or that.

But I never work on my cars any more. When I look under the hood now, all I see is a complex array of computers and microchips—nothing a person can get a wrench around. Yet the car I drive today, provided everything is working, is much more user-friendly than my old Volkswagen. The flow of gasoline to the cylinders, for example, is regulated by a small onboard computer, rather than by a clumsy mechanical carburetor.

A classic VW Beetle.

This personal story about cars turns out to be a pretty good allegory for the way in which the science of materials has developed in the twentieth century. In the beginning, industry turned out big, relatively simple things that were easy to understand and work with—iron wheels for railroads, steel springs for car suspensions, wooden chairs and tables for the home. Today, industry turns out items that perform the same jobs better, but that are made from new kinds of materials such as plastics, composites, and semiconductors. Instead of manipulating large chunks of material, we now control the way atoms fit together. Like modern cars, modern materials do their job well, but they cannot be made (or, usually, repaired) by a simple craftsperson working with simple tools.

So while the materials we use are becoming better at what they do and easier for us to use, it becomes harder and harder for us to understand what those materials are. I was able to fix my Volkswagen myself, but there is no way I can look under the hood of my present-day car and shift the atoms around in its microchip. In a sense, the improved performance of modern materials has been bought at the price of our ability to understand them. ●

Connection

Magnetic Data Storage

Computers rely on the ability to store information on magnetic media such as hard drives and compact discs. Remarkably, for the past several decades this capacity has doubled almost every year. This memory storage capacity (called "areal density") is measured in terms of the number of bits of information that could be stored on 1 square inch of material. (For historical reasons, the semiconductor industry does not always use metric units.) In mid-2001, the best commercially available storage was about 32 billion bits per square inch, and by mid-2002 capacity exceeded 50 billion bits per square inch. Experts predict that magnetic storage capabilities will surpass 1 trillion bits per square inch sometime in 2006.

Here's an amazing fact: according to some calculations, the total amount of information stored in all forms (cuneiform tablets, manuscripts, books, floppy disks, etc.) since the beginning of time is about 2.5 quadrillion bits. Some scholars predict that by 2006 we will be adding this much to the total stored information every 2 months.

The basic principle behind magnetic storage is the same as that behind the cassette tape. Tiny grains of magnetic material are lined up in a gel, and then a pickup head gets the information back out by detecting the directions of these grains' magnetic fields. The recording and reading heads move over the disk in a way analogous to the way an old-fashioned record-player arm moved over an LP record. Today's devices employ heads that float several atomic widths above the disk. The closer together you can get head and storage surface, the better off you are. The disk itself is a high-tech piece of equipment consisting of several layers of different materials. The working part, where the magnetic grains are, is about 200 atoms deep.

To compress more and more data onto a disk, these grains have to get smaller and smaller (in some laboratory systems they approach the size of individual atoms!). This process is inherently limited, however: if the grains are too small, then ordinary movement associated with thermal motion can erase the information minutes after it is written. In any case, current technology is expected to achieve 100 billion bits per square inch in the near future. As we show in the example and problems at the end of this chapter, the implications of this level of data storage are truly staggering. ●

Connection

Computers and the Internet

The worldwide network of computers known as the Internet has changed the way most of the world communicates. You can send personal e-mail to friends or computer files to co-workers that get to their destination within seconds. Doctors can look up a patient's X-ray films or MRI scans from the hospital database and show them on the computer screen to discuss treatment; researchers can find copies of past articles about virtually any topic as background for their own work. With all this information available from your computer, just how does the Internet work?

The heart of the Internet is a network of supercomputers located in various places around the world (Figure 25-9). They store data and are linked by conventional cables, optical cables, and even radio links to communications satellites. To access the data on these computers, you subscribe to an Internet service provider (ISP), which offers you a starting place from which to navigate the Web. When you tell your computer where you want to go, the information goes to a router maintained by the ISP that interprets each Web address as a location on a particular computer in the Internet and connects you to that location.

Of course, you can also download files from the Web. Often you can do so most easily through a server computer operated by your school or company. This server is connected by high-speed lines to the main backbone computers of the Internet, so you can simply connect to the server and access or download the material from there. ●

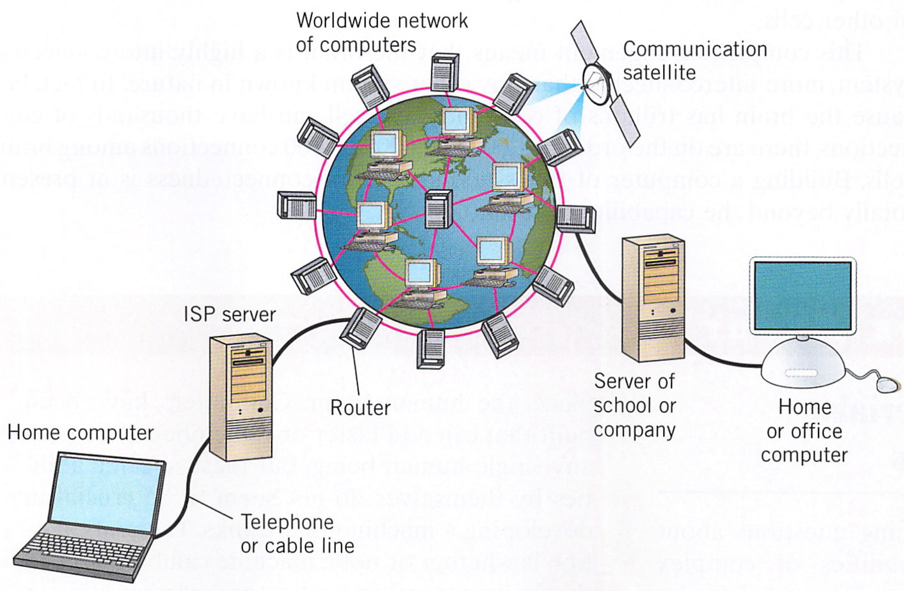


Figure 25-9. The Internet links numerous personal computers through a network of supercomputers located at various regions around the world.

Connection
The Computer and the Brain

When computers first came into public awareness, there was a general sense that we were building a machine that would in some way duplicate the human brain. Concepts such as *artificial intelligence* were sold (some would say oversold) on the basis of the idea that computers would soon be able to perform all those functions that we normally think of as being distinctly human. In fact, this scenario has not come to pass. The reason has to do with the difference between the basic unit of the computer, which is the transistor, and the basic unit of the brain, which is the nerve cell, or neuron.

The transmission of electric signals between the brain's neurons is fundamentally different from that between elements in ordinary electric circuits (see Chapter 18). This difference in signal transmission alone, however, does not make a brain different from a computer. A computer normally performs a sequential series of operations—that is, a group of transistors takes two numbers, adds them together, feeds that answer to another group of transistors that performs another manipulation, and so on. Some computers are now being designed and built that have some parallel capacity—machines in which, for example, addition and other manipulations are done at the same time rather than one after the other. Nevertheless, the natural configuration of computers is to have each transistor hooked to, at most, a couple of others.

Data, from *Star Trek: The Next Generation,* is an android. Computers that think like people have been a staple of science fiction stories, but could they really be built?

A nerve cell in the brain, however, operates in quite a different manner. Each of the brain's trillions of nerve cells connects to a thousand or more different neighboring nerve cells. Whether a nerve cell decides to fire—whether or not the signal moves out along the axon—depends in an unknown way on a complex integration of all the signals that come into that cell from thousands of other cells.

This complex arrangement means that the brain is a highly interconnected system, more interconnected than any other system known in nature. In fact, because the brain has trillions of cells and each cell can have thousands of connections, there are on the order of 1,000,000,000,000,000 connections among brain cells. Building a computer of this size and level of connectedness is at present totally beyond the capability of technology. ●

Thinking More About

Properties of Materials: Thinking Machines

One of the most intriguing questions about the ever-increasing abilities of complex computers is whether or not a computer can be built that could in some way mimic, or even replace, the human brain. Computers have been built that can add faster or remember more than any single human being, but these specific abilities by themselves do not seem to be crucial in developing a machine that thinks. The real question is whether or not a machine can be designed that is, by general consensus, regarded as alive or conscious.

World chess champion Gary Kasparov defeated the specially enhanced computer program Deep Blue in 1996. However, in 1997, Deep Blue defeated Kasparov in a rematch.

British mathematician Alan Turing (1912–1954) proposed a test to address this question. Called the "Turing test," it operates this way: A group of human beings sit in a room and interact with something through some kind of computer terminal. They might, for example, type questions into a keyboard and read answers on a screen. Alternatively, they could talk into a microphone and hear answers played back to them by some kind of voice synthesizer. These people are allowed to ask the hidden something any questions they like. At the end of the experiment, they have to decide whether they have been talking to a machine or a human being. If they can't tell the difference and the something is a machine, the machine is said to have passed the Turing test.

As of this date, no machine has passed the test (there have been occasional contests in Silicon Valley in which machines were put through their paces). But what if a machine did actually pass? Would that mean we had invented a truly intelligent machine? John Searle, a philosopher at the University of California at Berkeley, has recently challenged the whole idea of the Turing test as a way of telling if a machine can think by proposing a paradox he calls the "Chinese room."

The Chinese room works like this: An English-speaking person sits in a room and receives typed questions from a Chinese-speaking person in the adjacent room. The English-speaking person does not understand Chinese, but has a large manual of instructions. The manual might say, for example, that if a certain group of Chinese characters are received, then a second group of Chinese characters should be sent out. The English-speaking person could, at least in principle, pass the Turing test if the instructions were sufficiently detailed and complex. Obviously, however, the English speaker has no idea of what he or she is doing with the information that comes in or goes out. Thus, argues Searle, the mere fact that a machine passes the Turing test tells you nothing about whether it is aware of what it is doing.

Do you think a machine that can pass the Turing test must be aware of itself? Do you see any way around Searle's argument about the Chinese room? What moral and ethical problems might arise if human beings could indeed make a machine that everyone agreed has consciousness?

Summary

Doping with small amounts of another element modifies semiconductor material, usually silicon. Phosphorus doping adds a few mobile electrons to produce an *n*-type semiconductor, while aluminum doping provides positive holes in *p*-type semiconductors. Devices formed by juxtaposing *n*- and *p*-type semiconductors act as switches and valves for electricity. A **diode** joins single pieces of *n*- and *p*-type material, for example, to act as a one-way valve for current flow. **Transistors,** which incorporate a *pnp* or *npn* semiconductor sandwich, act as amplifiers or switches for current. **Microchips** can combine up to millions of *n* and *p* regions in a single integrated circuit.

Semiconductor technology has revolutionized the storage and use of information. Any information can be reduced to a series of simple yes-or-no questions, or **bits.** Eight-bit words, called **bytes,** are the basic unit of **digital** information used by most modern **computers.**

Key Terms

binary Presenting information as a string of 0s and 1s, representing off or on. (p. 540)

bit The smallest unit of information storage. (p. 540)

byte Eight bits of information (p. 540)

computer A machine that stores and manipulates information. (p. 542)

digital Presenting information in a numerical system that can have a fininte number of states. Binary is one form of digital information. (p. 540)

diode A semiconductor device formed from one *p*-type (positive charge carrier) and one *n*-type (negative charge carrier) region; typically used as a rectifier. (p. 533)

doping The addition of small amounts of an impurity to an element or compound to enhance its conduction properties. (p. 532)

microchip A complex array of *p*- and *n*-type semiconductors that constitutes a tiny integrated circuit. (p. 537)

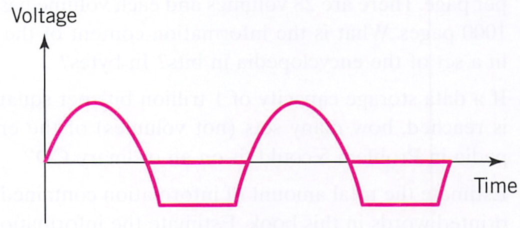
transistor An element in an electric circuit that regulates current or voltage flow. (p. 535)

Review

1. Semiconductors remain the anchor of the information revolution. What are some of the devices they are used in?
2. Identify five objects in your home that use semiconductors. What other kinds of materials with special electric properties are found in all of these five objects?
3. What common chemical element is used to make semiconductors? What aspect of this element's structure allows it to function as a semiconductor?
4. What is doping? What are the consequences of having an extra electron available from a doping agent?
5. Phosphorus is a common doping agent. What is it about the electron structure of phosphorus that makes it so useful as such an agent?
6. What does the term "n-type semiconductor" stand for? Why is a semiconductor doped with phosphorus called an *n-type semiconductor*?
7. Why is a semiconductor doped with aluminum called a "p-type semiconductor?" How does this type of semiconductor work?
8. What happens when an *n*-type and a *p*-type semiconductor are placed back to back? What happens to the charges of each? What is a device that works this way called?
9. In which direction will a permanent electric field push an electron in a diode?
10. How does a diode act as a one-way gate?
11. What is a rectifier, and how does it convert the AC current that enters your home into the DC current used by many household appliances?
12. What is a photovoltaic cell?
13. How does a photovoltaic cell generate electric current? When is the use of such a cell most desirable and efficient?
14. What is the difference between base-load electricity requirements and peak-load requirements? How does this affect the feasibility of solar power?
15. How is a modern transistor designed and what is its function? Describe the importance of this device to the information technology industry.
16. What is a microchip? Why is this object also called an integrated circuit?
17. Compare the size of the original microchips to their size today.
18. Describe a process by which a microchip is made. How can a *p*-type chip be made?
19. How can you amplify a weak electric current with a transistor?
20. How can a transistor act as a switch?
21. What is the unit of information that scientists typically use? What does this represent, and why is this so useful for information devices that rely on electrical circuitry? In other words, how can such a unit of information be expressed using electricity?
22. What is the information content of a single letter of the alphabet in bits?
23. How many bits equal a byte?
24. How can sound wave pressure be converted into digital form?
25. Describe how a digital signal is converted back to sound in a CD player.
26. How does a computer store information? What is the connection between tiny magnetic particles and a bit of information?
27. What is the role of the CPU in a computer?
28. What is nanotechnology, and how has it helped increase the capacity of information systems?
29. Compare a transistor with the nerve cells in your brain. What implications do the differences between these two have for the development of artificial intelligence, if any?
30. What is the Turing test? What does it tell us about the ability of a machine to think?

Questions

1. If water in a pipe is analogous to electricity in a wire, what plumbing equipment is analogous to a diode? To a transistor? A water storage tank is analogous to which electrical device?
2. Would a silicon semiconductor doped with boron be *n*- or *p*-type? How about one doped with arsenic? (*Hint:* Look at the periodic table.)
3. Does the Turing test seem like a reasonable way to judge whether a computer has consciousness? Explain.
4. In order to make an *n*-type semiconductor, silicon can be doped with a small amount of phosphorus. Why is arsenic also a good element to use as a dopant? (*Hint:* see the periodic table of elements.)
5. Compare the electrical conductivity of two phosphorus-doped silicon semiconductors where one has twice the amount of phosphorus as the other.
6. Would a diode result from taking two *n*-type semiconductors and placing them together?
7. The figure represents a graph of voltage versus time for an electric power source. Make a graph of current versus time if this source is connected to a diode rectifier. Make another graph of current versus time if the diode is reversed.



8. Which would be more effective at making a photovoltaic cell work, infrared light or ultraviolet light? Explain.
9. Which method most likely requires more digital storage capacity, storing the words of a song using a word-processing program or storing an actual recording of the song on a CD? Explain.
10. You have five pennies and five nickels in a hat. You draw them out randomly and flip each coin. How many bits of information are required to record the sequence of coins drawn from the hat (e.g., penny, nickel, nickel, penny, . . .)? How many bits are required to record the sequence of coins and the sequence of heads/tails?

Problem-Solving Example

EXAMPLE 25-1

Recording Your Life

Suppose that a person has a miniature camera implanted at birth and that every 10 seconds the camera takes a picture of what that person sees. Suppose that the picture is in black and white and is somewhat grainier than that on commercial television, so that its information content is 1 Mbit. If information storage technology reaches 1 trillion bits per square inch, how much of that person's life could be recorded on a disk that is 12 inches across?

SOLUTION: If information is coming in at the rate of 1 Mbit every 10 seconds, then the total information generated in 1 day is

$$1 \text{ Mbit} \times 6 \text{ pictures per minute} \times 60 \text{ minutes per hour} \times 24 \text{ hours per day} = 8640 \text{ Mbits} = 8.64 \text{ Gbits}$$

A 12-inch disk has a radius of 6 inches and an area of

$$A = \pi r^2 = 3.14 \times 36 = 113 \text{ in}^2$$

At 1 trillion bits per square inch, the disk can hold an amount of information I equal to

$$I = 113 \text{ in}^2 \times 10^{12} \text{ bits/in}^2 = 1.13 \times 10^{14} \text{ bits}$$

Thus, the total number of days of information that can be stored on the disk is

$$\frac{1.13 \times 10^{14} \text{ bits}}{8.64 \times 10^9 \text{ bits/day}} = 13{,}078 \text{ days}$$

which is 35 years. Thus, at this rate, a person's entire life could be recorded on two 12-inch disks! ●

Problems

1. There is an effort in the world today to convert television into so-called high-definition TV (HDTV). In HDTV, the picture is split up into as many as 1100 by 1100 (as opposed to 525 by 525) pixels. What is the information content of an HDTV picture? What is the information content that must be transmitted each second in an HDTV broadcast?
2. Construct a set of yes-or-no questions to specify any letter of the alphabet, both upper- and lowercase, and all digits from 0 to 9.
3. Construct a set of yes-or-no questions to specify any state in the United States.

4. How many seconds do you have to listen to a CD to receive as much information as is contained in an average book?
5. The *Encyclopedia Britannica* contains about 1800 words per page. There are 28 volumes and each volume has about 1000 pages. What is the information content of the words in a set of the encyclopedia in bits? In bytes?
6. If a data storage capacity of 1 trillion bits per square inch is reached, how many sets (not volumes) of the encyclopedia in Problem 5 could fit on an ordinary CD?
7. Estimate the total amount of information contained in the printed words in this book. Estimate the information content of the illustrations in this book.
8. Rosa writes a 20-page paper for her extra-credit history grade. Each page has an average of 26 lines, with 12 words per line.
   a. How many bits of information has Rosa generated in her paper?
   b. Can she store her paper on a normal-mode 3.5-inch floppy disk?
9. The Cyrillic alphabet (used to write Russian and some other Eastern European languages) was devised in the ninth century and had 43 letters, whereas the alphabet used for modern Russian has 30. How many bits would it take to specify a single letter in each of these alphabets?
10. The version of modern written Japanese called "kanji" has 1945 different characters. How many bits would it take to specify a specific kanji character? Compare this to the number required in languages that have alphabets.
11. Show that a 16-bit number can represent any of 64,000 states, as it does in a CD recording.

Investigations

1. Some applications of information technology present real ethical challenges. For example, 24-hour video surveillance in public places is being increasingly used, both for simple security and, for example, to ticket a driver who runs a red light. The ultimate goal of some security agencies is for computerized pattern recognition technology to accurately match an image on a camera to an image stored in a photographic database, allowing police to uniquely identify wanted criminals and apprehend them. How does pattern recognition technology work and how accurate is it currently? What are the potential benefits and what are the potential pitfalls? How do you feel about having such technology in use everywhere? How prevalent is its use now? Considering the technological feasibility, the ethics, and the politics of it, how widespread do you think the use of this technology will be within 10 years?
2. The Internet is one of the most widely recognized applications of information technology. How does it affect your life daily and, especially, how does it affect how you socialize and how you gather information? Do you trust the information found on it? What are your standards of proof with electronic information? How do you yourself verify facts with it? From a social standpoint, would your interactions with other people be at all different without the Internet and its associated e-mail and chat rooms that allow quick online communication? Do you think this technology drastically improves your life or just makes it different?
3. What is DNA computing? How does it work? What are the potential advantages and disadvantages of this type of technology?
4. In this chapter the issue of whether a machine could ever think was raised; the Turing test was mentioned as one test for how we might probe this question and the Chinese room was offered as a criticism of this test. Investigate the history of the search for thinking machines. What are the criteria for a machine that actually thinks? How did human intelligence and the brain evolve over time? Could a machine ever duplicate this process? Summarize the arguments for and against the proposition that it will eventually be possible for machines to think.
5. The ENIAC (Electronic Numerical Integrator and Computer), built in the 1940s, was one of the first all-electronic digital computers. Investigate the ENIAC and compare it to an ordinary desktop PC in use today. How powerful is a PC today compared to the ENIAC? How large is each? In terms of materials that are involved, what was used instead of the typical semiconductors and microchips in use today? What kind of maintenance did the ENIAC require, and how was information stored?

WWW Resources

See the *Physics Matters* home page at **www.wiley.com/college/trefil** for valuable web links.

1. **http://www.pbs.org/transistor/** A historical/tutorial online exhibit accompanying the PBS program, *Transistorized!*, co-produced with the American Institute of Physics.

2. **http://www.101science.com/transistor.htm** A transistor tutorial, containing simulations describing typical circuits, applications, manufacturers, and standards.
3. **http://korea.park.org/Japan/NTT/MUSEUM/html_st/ST_menu_4_e.html** *The Basics of Electronic Communication,* a cartoon-illustrated tutorial by the NTT Digital Museum.
4. **http://www.computer50.org/** The University of Manchester site dedicated to the 50th anniversary of the birth of the modern computer.
5. **http://www.computer-museum.org/exhibits/pccomeshome/index.html** *The Computer Comes Home: A History of Personal Computing.* Online exhibit of the Computer Museum of America.
6. **http://micro.magnet.fsu.edu/electromag/java/cd/index.html** *How a CD Works,* from Florida State University.
7. **http://micro.magnet.fsu.edu/electromag/java/harddrive/index.html** *How A Hard Drive Works,* another short and simple tutorial from Florida State University.

26 The Nucleus of the Atom

KEY IDEA

The mass of an atom is concentrated in its nucleus, which is held together by powerful forces.

PHYSICS AROUND US . . . Enjoying Empty Spaces

It's great to be lying on the beach, lulled by the sound of the surf, soaking up the Sun. There's a great feeling of peacefulness in the wide-open spaces, with no crowding of people or buildings, just wind and air and water. Away from the pressures of school and work, time seems to stand still.

In such a relaxing setting, it's hard to imagine that your own body is also mostly empty space. Every atom has electrons whizzing around with various amounts of energy, but inside the cloud of electrons is almost nothing but empty space. Yet at the very center of each atom, in a tiny volume only about one-hundred-thousandth of the size of the electron cloud, is the nucleus that holds almost all the mass of the atom.

Most of the time, the nucleus stays in the middle of the atom and nothing happens. Electrons collide and change energy, atoms bond with other atoms, and the nuclei just go along for the ride, never changing. But some nuclei do change, and they emit highly energetic particles. In fact, thousands of energetic particles are passing unnoticed through your body every second. Some of those speeding particles are damaging your cells, breaking apart bonds in the molecules that control critical functions of metabolism and cell division.

But don't lose a moment worrying about this ubiquitous background radioactivity. Since the dawn of life, low levels of radioactivity in rocks, soils, oceans, and air have bathed every living thing. This radioactivity is a natural part of the everyday environment on Earth, like the ocean and beach, and our bodies have evolved mechanisms to repair the damage it causes. But radioactivity also reveals much about the inner structure of the atom.