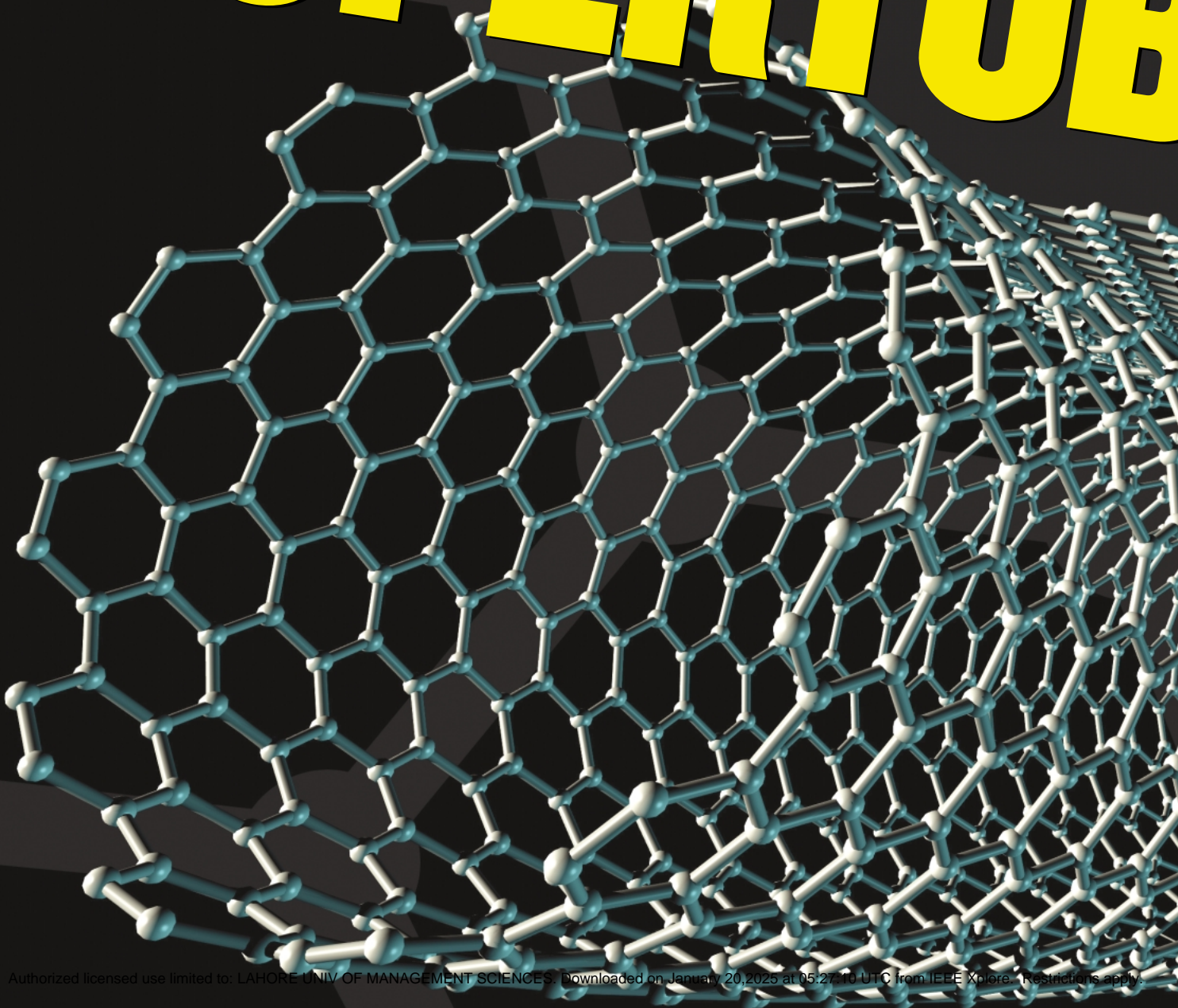


The unique properties of carbon nanotubes
may make them the natural successor to
silicon microelectronics **BY PHAEDON AVOURIS**

SUPERTUBE



BES

TOTALLY TUBULAR: The atoms in carbon nanotubes occupy the corners of hexagons arranged in a honeycomb pattern similar to that in a sheet of graphite [center]. Rolled up, the graphite layer forms a tube that can be either metallic [bottom] or semiconducting [top].

ILLUSTRATION: JOHN MACNEILL



arbon nanotubes: they can be far stronger than steel, lighter than aluminum, and more conductive than copper. Their potential applications range from ultrathin, breathable, waterproof fabrics to bright, rugged flat-panel displays for televisions and computer monitors.

Carbon nanotubes will undoubtedly be the wonder material of the 21st century. And though their list of potential applications is long and

dazzling, none are more important than those envisioned in electronics. Already, researchers have built a variety of carbon-nanotube electronic and optoelectronic devices: transistors, diodes, light emitters, and detectors. Most remarkable, you can get all those different functions with a single device, merely by altering the voltages you apply to it.

In fact, so attractive are carbon nanotubes' electrical properties that researchers are already eyeing them as replacements for silicon circuits. Since individual nanotubes can be created to be metallic conductors or semiconductors, you can use some of them as transistors and others as the connections between transistors—the two main ingredients of an integrated circuit. And that's very good news to electronics researchers who are approaching fundamental physical limits as they strive to scale conventional CMOS silicon circuits down into a realm in which certain transistor parts are only tens of atoms wide.

The steady reduction in the dimensions of transistors on ICs has been the main force behind the regular leaps in the level of performance of silicon ICs over the past four decades. However, no one expects those leaps to go on forever. In fact, technologists expect that those physical limits will become a serious problem within a decade. Of course, huge industries have grown accustomed to regular increases in computing power and memory-chip density, and they are already preparing for their long-term future by investigating potential successors to ordinary CMOS ICs.

Carbon nanotubes are one of the most promising of the technologies that might someday pick up where conventional CMOS devices leave off. An electronics industry based on nanotubes could preserve a lot of what's good about existing silicon technology—

the logic circuits and much of the manufacturing process—but base it on new materials that get around the majority of problems that would probably doom any attempts to make extremely small CMOS devices. So groups all over the world are making and investigating nanotube devices. My group, at the IBM Thomas J. Watson Research Center in Yorktown Heights, N.Y., is evaluating the potential of carbon nanotubes to augment and maybe ultimately replace today's ubiquitous silicon CMOS.

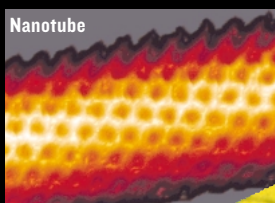
TO UNDERSTAND CARBON NANOTUBES, you first have to understand carbon. It exists in two crystalline forms: graphite and diamond. Nanotubes are structurally similar to graphite, the chief ingredient in pencil lead.

The first person to see carbon nanotubes was Sumio Iijima of NEC Corp. in Tokyo, who discovered them in 1991 while studying electron microscope images of the soot produced by electrical discharges between carbon electrodes. He saw molecules made up of carbon atoms, cylindrical in shape, exquisitely thin and impressively long. These early structures had the form of cylinders within cylinders, like Russian *matryoshka* nesting dolls. Then, in 1993, Iijima and Donald Bethune of IBM independently found that adding small amounts of metal catalysts to the carbon electrodes could produce nanotubes that were not nested together; that is, each nanotube was one macromolecule made of a single wall of carbon atoms. This achievement was significant, because nanotube transistors and circuits use such single-walled nanotubes.

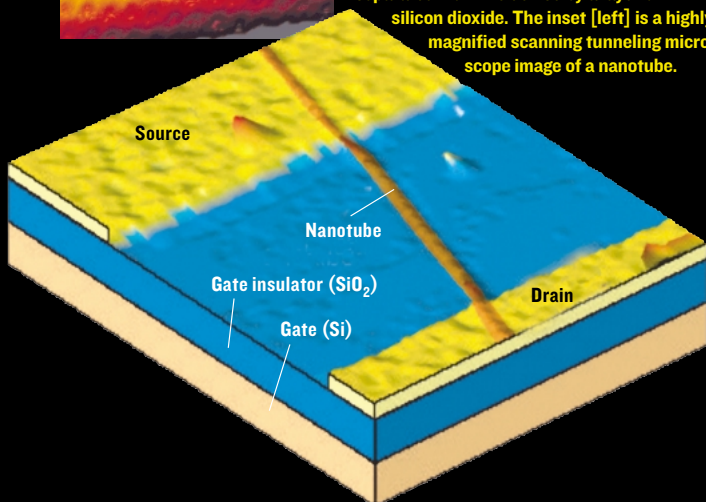
In graphite, the carbon atoms are arranged into hexagons that form a honeycomb pattern [see illustration, "Totally Tubular"]. A nanotube can be viewed as a single layer of graphite rolled into a seamless cylinder. One of the most alluring features of nanotubes as electronic devices is that you can change a device's characteristics merely by altering the physical traits of the nanotube. Two key traits are the width of the graphite layer that is rolled to make the tube, which determines the nanotube diameter, and the orientation of the honeycomb pattern with respect to the nanotube axis. In some nanotubes, the honeycomb pattern lines up with the nanotube axis; in others it spirals around the axis like the stripes on a candy cane. The combination of diameter and twist determines whether the nanotube is metallic or semiconducting.

In semiconducting nanotubes, the diameter of the tube affects how much energy an electron needs to move from the valence band, where it is bound to an atom, into the conduction band, where it is free to move about the semiconductor and conduct electricity. This required energy is called the band gap of the material. The importance of the tube diameter in determining the band gap comes from a quantum-mechanical property of electrons: they are not simply small, charged particles but have wave properties as well. And just as with light waves, the electron's wavelength determines its energy: the shorter the wavelength, the higher the energy.

Because of its wave nature, an electron can experience interference, just as light waves or sound waves do. As a result, an electron wave moving around the circumference of a nanotube can be only in states in which an integral number of wavelengths can fit around the nanotube circumference. Otherwise, the electron will interfere destructively with itself. This requirement restricts the electron's energy to certain discrete values (energy states). The smaller the nanotube diameter, the larger the separation between these allowed energies. An electron jumping from the valence band into the conduction band must have enough energy to jump into at least the lowest of these conductive energy states. That, in turn, determines the band gap.



AN EARLY PROTOTYPE: The surface in this illustration is an image from an atomic force microscope. A single nanotube is the transistor channel between the two gold electrodes that form the source and drain. The gate is the heavily doped silicon substrate, separated from the device by a layer of silicon dioxide. The inset [left] is a highly magnified scanning tunneling microscope image of a nanotube.



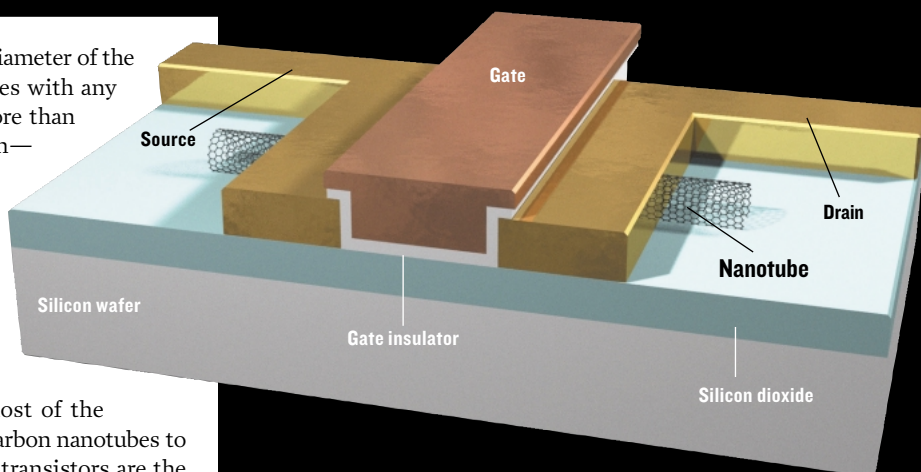
The upshot is that by changing the diameter of the nanotubes, researchers can produce devices with any band gap from 0 (a metallic nanotube) to more than 1 electronvolt—roughly the band gap of silicon—and all gap values in between. This feature allows us to make devices that turn on and off at different voltages, which we can tailor for different applications. That kind of versatility isn't possible with conventional devices, which are limited to the band gap of whatever semiconductor they are made of.

MUCH OF OUR WORK AT IBM, and indeed most of the efforts of our colleagues, is devoted to using carbon nanotubes to make field-effect transistors. After all, these transistors are the building blocks of CMOS ICs. To understand why nanotubes are promising candidates for making field-effect transistors, consider some of the factors that ultimately limit the miniaturization of conventional field-effect devices. The silicon field-effect transistor is a switch with three terminals: source, drain, and gate. A voltage applied to the gate acts as a spigot that controls the current between the source and the drain. The current moves through the channel—a thin layer of silicon under the gate, electrically insulated from it—that connects the source and the drain.

To make a transistor smaller and thus enhance its performance, you have to simultaneously diminish multiple structural features and change the operating parameters of the device. For example, moving the source and drain 30 percent closer together requires that several other characteristics also decrease by about 30 percent. These include the power supply voltage, the thickness of the insulation between the gate and the channel, and the widths of the wires that connect different transistors.

Unfortunately, orchestrating all those reductions isn't so easy nowadays. Consider the gate insulation. In a state-of-the-art IC with a channel length of about 100 nanometers, the gate insulation is just 1.5 nm thick. That insulation is supposed to prevent electrons from leaking between the gate and the channel. But when the insulation gets too thin, electrons can pass right through it, thanks to a quantum-mechanical phenomenon called tunneling. The same phenomenon limits how close together you can bring the source and drain—that is, how short the channel length can be. The unwanted flow of electrons due to tunneling undermines the function of the transistor as a switch and adds to the power density, already soaring on complex ICs, which pushes heat to unmanageable levels. Then, too, reducing the width of the wiring that connects the transistors into circuits increases their resistance, slows circuit-switching speeds, and hastens disintegration of the metallic wires as a result of the high current density.

Consider, on the other hand, a carbon-nanotube field-effect transistor. Researchers created the first crude devices just six years ago in separate projects at Delft University in the Netherlands and in our laboratory at IBM. To make these first prototypes, experimenters covered a silicon wafer with a thick silicon-dioxide film and then fabricated gold or platinum electrodes on it using standard semiconductor manufacturing techniques [see illustration, "An Early Prototype"]. A single carbon nanotube was then positioned as a bridge between two electrodes. The electrodes became the source and drain of the transistor, and the nanotube played the role of the channel. The underlying silicon wafer, heavily doped with impurities to make it a good conductor, served



THE SECOND GENERATION: More advanced nanotube transistors have individual gates on top of the device, separated from the nanotube channel by a thin layer of silicon dioxide. Metal electrodes form the source and drain.

as the gate electrode. Applying the appropriate voltage to it turned the current through the nanotube on or off.

Initially, my colleagues and I thought that the mechanism for switching the current on and off in the nanotube transistor was exactly the same as that for a silicon field-effect transistor (in which the role of voltage on the gate is simply to modify the conducting properties of the bulk channel). Further experiments, however, showed some important differences.

IN THE EXPERIMENTS, we sought to optimize the carbon-nanotube devices by shrinking their dimensions in the same manner that was responsible for the improvements in silicon transistors; that is, we proportionally reduced channel lengths, gate voltages, and gate insulation thicknesses. Although we saw changes in the nanotube properties, the changes did not agree quantitatively with our expectations based on the scaling of silicon transistors.

Through more experiments and theoretical modeling, we realized that in the nanotube device, the transistor action occurs at the contact points between the metal electrodes and the carbon nanotube, where the contact between the two dissimilar materials sets up an energy barrier that prevents electrons from crossing between the metal electrode and the semiconducting nanotube. Increasing the voltage on the gate thins the barrier and turns the nanotube transistor on.

The barrier is similar to one that forms in a conventional diode, which lets electrons easily flow in one direction only. Connecting an ordinary semiconductor to a metal forms a Schottky diode, named for Swiss-German physicist Walter Schottky, who discovered the effect. In the case of the nanotube, the energy barrier set up between the metal electrode and the carbon nanotube is a Schottky barrier, except that the semiconductor—the carbon nanotube—is one-dimensional, not three-dimensional. A 1-D barrier tends to be much narrower, and under the influence of the nanotube's gate voltage, the barrier becomes so thin that electrons can pass through it, thanks to tunneling.

Our first nanotube transistors were simple switches with current-on to current-off ratios of about 100 000, compared with 100 000 to a million for a typical silicon transistor. But they had weak currents and high resistance between the metal contacts and the nanotubes. We therefore concentrated on improving the metal-to-nanotube contacts and enhancing the coupling between the gate and the nanotube. At the same

time, we worked on methods of wiring together individual transistors to form logic circuits, and we also fabricated novel circuits by building several transistors along the length of a single nanotube—a procedure that could conceivably simplify the manufacture of nanotube ICs.

In the early nanotube transistors, the current carriers were holes—atoms in the valence band that are missing an electron and therefore act like positively charged electrons. But the benefit of CMOS technology—lower power consumption—shows the advantage of using two different kinds of transistors, one that works by conducting with holes (p-type) and the other with electrons (n-type). At the beginning of this decade, researchers in our lab and elsewhere produced the first n-type nanotube transistors by adding to the p-type devices foreign atoms that donate electrons—a process called doping—and then building CMOS-type logic gates using both p-type and n-type nanotube transistors integrated together on the same IC.

By 2001, our carbon-nanotube field-effect transistors were superior to state-of-the-art silicon devices, as gauged by their output current and transconductance (the measure of how much the current changes with a change in the gate voltage). Instead of using the silicon wafer as the gate, these second-generation devices had individual gates on top of the transistors, separated from the nanotubes by thin silicon dioxide films or high-dielectric-constant materials [see illustration, “The Second Generation”]. Since then, we and other teams have reduced the contact resistance significantly by combining different metals and coming up with new ways to fabricate the metal-to-nanotube contacts. Our latest devices involve double-gate transistor structures that let us convert Schottky barrier nanotube transistors to bulk-switched transistors operating exactly like silicon MOSFETs (metal-oxide semiconductor field-effect transistors). We have also achieved the same effect by using structures that involve selective doping of the metal-to-nanotube contacts. Other efforts involve the measurement and improvement of the ac characteristics of nanotube transistors by minimizing their parasitic capacitance and resistance.

THE UNIQUE ELECTRICAL PROPERTIES of carbon nanotubes are to a large extent derived from their 1-D character and the peculiar electronic structure of graphite. Take, for instance, their extremely low electrical resistance. Resistance occurs when an electron collides with some defect in the crystal structure of the

material through which it is passing. The defect could be an impurity atom, a defect in the crystal structure, or an atom vibrating about its position in the crystal. Such collisions deflect the electron from its path.

But the electrons inside a carbon nanotube are not so easily scattered. Because of their very small diameter and huge ratio of length to diameter—a ratio that can be up in the millions or even higher—nanotubes are essentially 1-D systems, as noted before. In a 3-D conductor, electrons have plenty of opportunity to scatter, since they can do so at any angle. Any scattering gives rise to electrical resistance. In a 1-D conductor, however, electrons can travel only forward or backward. Under these circumstances, only backscattering (the change in electron motion from forward to backward) can lead to electrical resistance. But backscattering requires very strong collisions and is thus less likely to happen. So the electrons, having far fewer possibilities to scatter, travel for long distances before they encounter an obstacle that can send them careening in the opposite direction.

This reduced scattering gives carbon nanotubes their very low resistance. Electrical transport in high-quality metallic nanotubes is ballistic. That is, the electrons can travel a few micrometers without a collision, even at room temperature. By comparison, electrons in an excellent conductor like copper will get only about 40 nm at most before they scatter. Semiconducting nanotubes biased at low voltages are also ballistic on a scale of a few hundred nanometers, more than is needed to fabricate nanotube-based field-effect transistors.

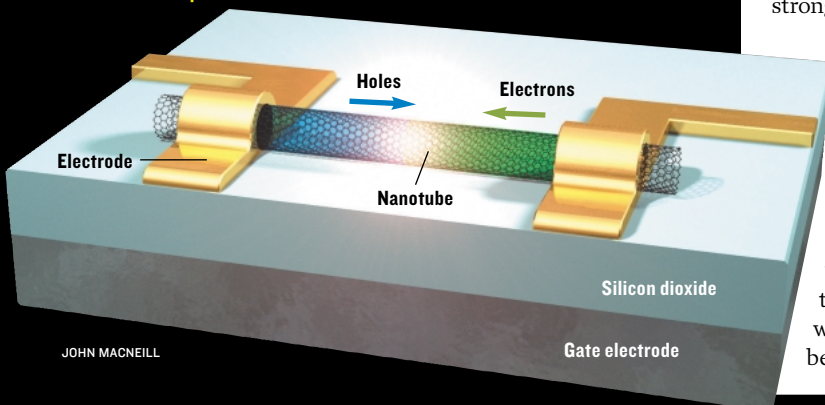
As a result of their low resistance, the energy dissipated in carbon nanotubes is very small, alleviating the issue of dissipated power density that plagues silicon circuits. In addition, metallic nanotubes can carry 100 to 1000 times as much current per given cross-sectional area as common metals like aluminum or copper.

Another advantage of a carbon nanotube is due to the character of its surface. Atoms vary in the number of bonds they can make, which for silicon is four. So in the interior of a silicon crystal, each atom bonds with four nearby silicon atoms. But at the surface, some bonds aren’t connected to other atoms and can trap stray electrons. Such charged sites can destroy transistor action. Chip makers, therefore, expose the silicon surface to oxygen atoms, so that they form bonds with the silicon atoms at the surface and build a silicon dioxide film that also serves as the transistor’s gate insulation.

By contrast, the carbon atoms in nanotubes do not have any leftover bonds, so there is no need to grow a film on the surface in order to tie up the free bonds and, more important, no need to restrict the gate insulator to silicon dioxide. This fact opens the door for the use of other, superior materials (ones with higher dielectric constants) to insulate the transistor’s gate. These insulators can be physically thicker yet still allow the gate to exert strong control over the channel. The greater thickness eliminates tunneling, while the strong concentration of the electric field allows the gate to control the channel. Freed from the problematic effects of thinner and thinner silicon dioxide gate-insulating films, the resulting device is a faster, smaller nanotube-based field-effect transistor.

FIELD-EFFECT TRANSISTORS made with carbon nanotubes rather than silicon have other advantages as well. Silicon transistors have a doped channel. As they are scaled down, the doping of the channel has to increase proportionately, while its volume decreases. The fluctuations in the number of dopants in this small volume from transistor to

AND THEN THERE WAS LIGHT: A very useful property of nanotube transistors is their ability to emit light. Under the right biasing conditions, electrons and holes enter the nanotube channel from opposite ends of the device and give off light or heat when they meet. This property raises the intriguing possibility of combining electronics and photonics into circuits built with the same materials.



transistor produce important differences in their switching properties and degrade the overall performance of the system. Similar degradation of performance occurs because of variations in other parameters (such as the channel length), which increase with decreasing size. In contrast, nanotube transistors can operate even without dopants and are less sensitive to differences in channel length. Rather, they depend on the diameter of the tube and the degree of twist of the honeycomb pattern. And both of these qualities are determined by the chemical synthesis of the nanotubes.

I should point out that carbon nanotubes do not necessarily allow the fabrication of devices with channel lengths of just a few nanometers, because limitations such as those imposed by tunneling from source to drain will afflict nanotube transistors just as they do silicon devices. Moreover, the lengths of nanotube devices are still likely to depend on the ability of lithographic tools to print small structures on wafers. Nevertheless, they will allow us to build, with relative ease, devices with superior performance than that of ultrasmall silicon devices.

THE ABILITY TO EMIT and detect light is another advantage of nanotubes. This feature raises the possibility of a future optoelectronics technology based on carbon nanotubes.

When the thickness of the gate insulator of a nanotube transistor gets sufficiently small, the transistor becomes ambipolar—that is, electrons conduct the current when the gate voltage is positive, and holes conduct it when the gate voltage is negative. Such ambipolarity is undesirable in electronic applications, and we have shown that it can be eliminated through proper design of the gate. But it is very valuable in optoelectronic applications.

Because of the ambipolar nature of a nanotube, under appropriate bias conditions, electrons and holes can enter the channel simultaneously from the opposite ends of the device. And when they meet, they can release energy in the form of heat or light. To be more precise, what is really happening is that electrons fall from the conduction band to the valence band, releasing the band-gap energy in the process [see illustration, “And Then There Was Light”].

In 2003 my team at IBM produced the first single-molecule, electrically controlled light source. Unlike ordinary light-emitting diodes, this carbon-nanotube light source does not rely on dopants to create light. Moreover, a diode is a two-terminal device, while the nanotube light source has three terminals. It is, in effect, a new addition to the electronics bestiary: a light-emitting transistor. Using its third terminal, we can control not only the intensity of the emitted light but also the position of the emitting spot along the length of the nanotube [see image, “Controlling Light”].

The energy that the electron gives off when it falls from the conduction band into the valence band—the band-gap energy—determines the wavelength of the emitted light.

And as we discussed above, it is the diameter of the nanotube that determines the band gap. So we can make

nanotube light sources with different wavelengths by using nanotubes with different diameters. We have also been able to perform the reverse process: to generate an electrical current (and voltage) by exposing the carbon-nanotube transistor to light. Thus, we have both a molecular light source and a light detector.

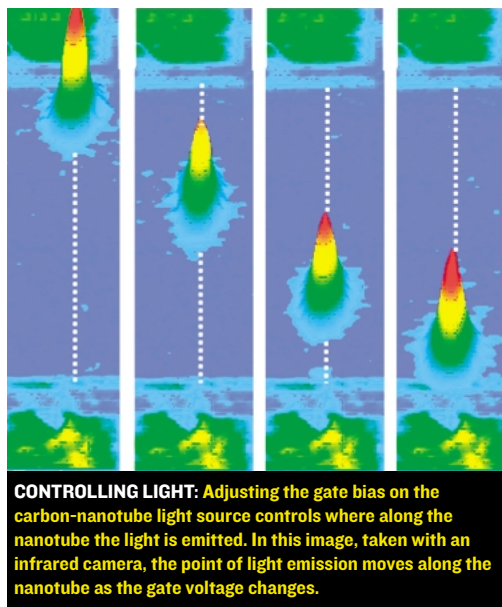
Despite the spectacular properties of carbon nanotubes, we will have to overcome many serious hurdles before we can create an electronic nanotechnology based on them. First, we need an approach for making them that leads to a homogeneous material. Current techniques produce a mixture of semiconducting and metallic nanotubes with different diameters and different amounts of twist in their structures. If we are to make ICs out of nanotubes, we must be able to control completely the nature of the nanotubes we create.

Recently, there has been significant progress in the area of selective synthesis—the process of making nanotubes with specific diameters and twists. Samples containing only a small number of different nanotubes with similar diameters have been made by proper selection of catalysts, starting materials, and reaction conditions. At the same time, researchers are also coming up with chemical and physical techniques for separating the different types of nanotubes after they have been made.

The next important step will be the integration of carbon-nanotube devices into complex CMOS-type circuits, because they will become the building blocks of a new generation of nanotube ICs. This effort will likely employ a mixture of techniques. We may apply lithographic techniques similar to those used for patterning silicon ICs to define the

overall structure of the chip. To build the transistors themselves, researchers are working on ways to let the nanotubes assemble themselves in just the right configurations. These self-assembly techniques are still in the early stages of development. Then, too, taking into account carbon nanotubes’ unique properties may lead to new types of circuits that are faster and smaller than the standard circuits used in ICs today.

As the dimensions of silicon CMOS transistors continue to shrink well into the next decade, problems resulting from increasing power dissipation, leakage currents, and variations in device parameters will continue to rise. If all goes well, carbon-nanotube electronics will be poised to take over before the problems encountered by the continual downscaling of silicon CMOS dimensions become insurmountable. ■



CONTROLLING LIGHT: Adjusting the gate bias on the carbon-nanotube light source controls where along the nanotube the light is emitted. In this image, taken with an infrared camera, the point of light emission moves along the nanotube as the gate voltage changes.

TO PROBE FURTHER:

A general reference on carbon nanotubes is *Carbon Nanotubes: Synthesis, Structure, Properties and Applications*, eds. M.S. Dresselhaus, G. Dresselhaus, and Ph. Avouris (Springer-Verlag, Berlin), 2001.

“Carbon nanotube electronics,” by Ph. Avouris, J. Appenzeller, R. Martel, and S. Wind, *Proceedings of the IEEE*, Vol. 91, 2003, pp. 1772–84, is a general discussion of the electrical properties of carbon nanotubes.