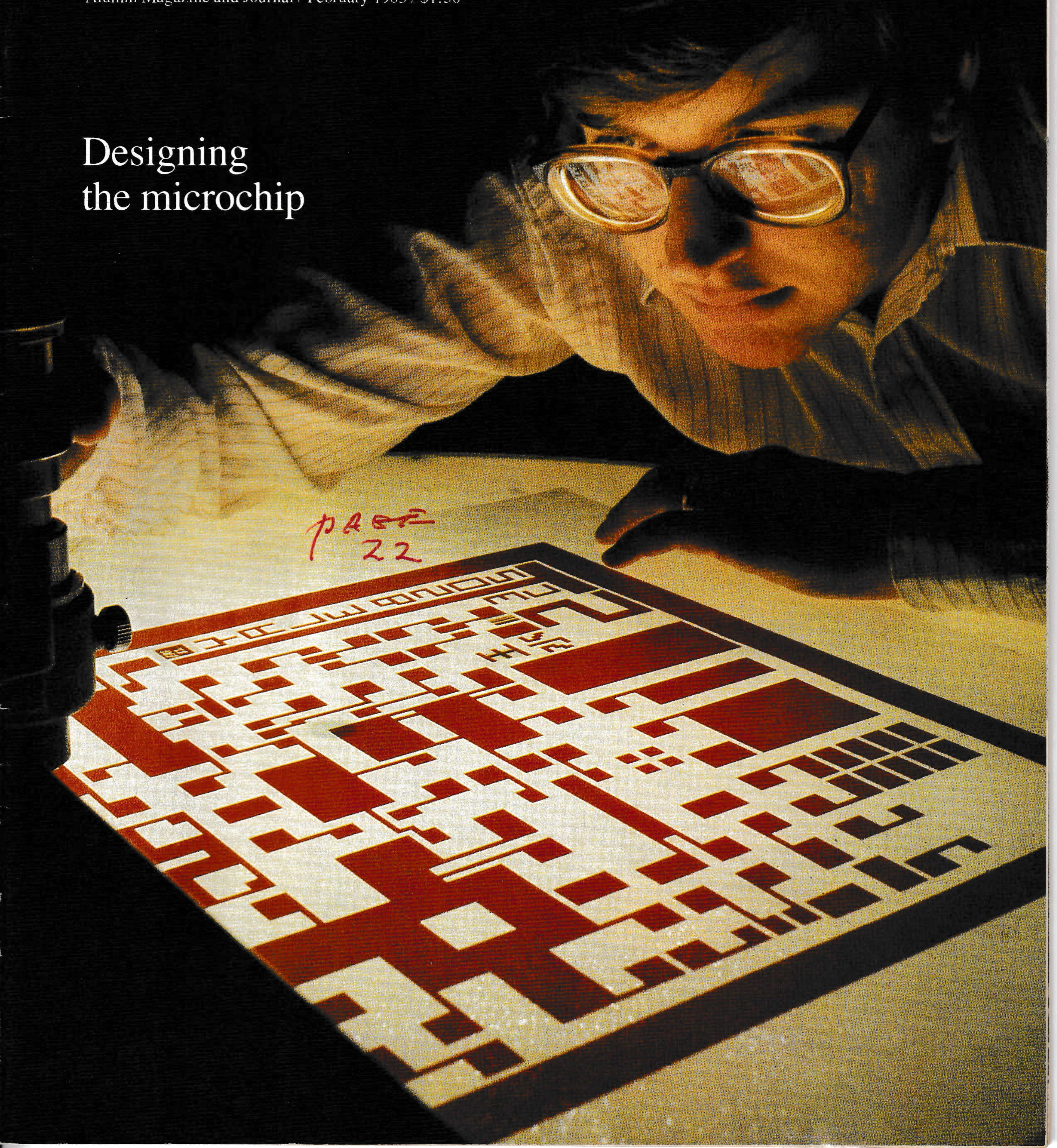
Yale

Alumni Magazine and Journal / February 1983 / $1.50

Designing the microchip



Small and beautiful, the potent microchip

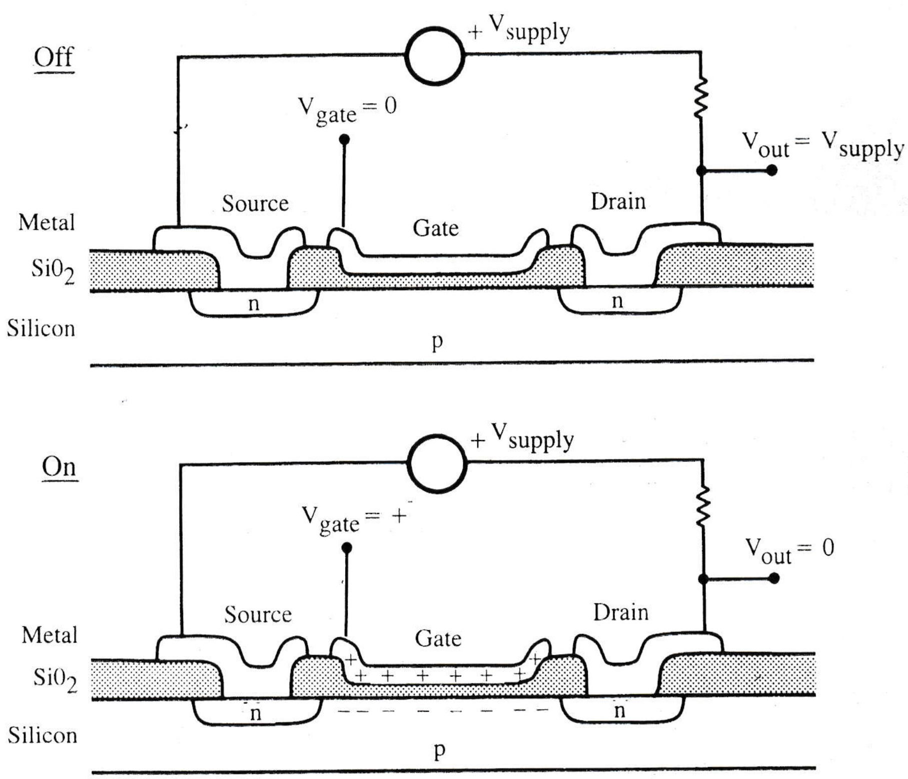
Yale engineering labs explore development and applications

by D. Kimball Smith

Scientists and philosophers were once wont to exercise their minds with questions of how many angels could dance on the head of a pin. It might have seemed a compelling problem at the time, but their answers tended to be imprecise. Today, the density of angels is no longer of premier importance, but in the field of microelectronics scientists and engineers are still concerned with filling tiny spaces with even tinier things. Instead of dancing, however, these microscopic devices are switching, computing, measuring and, in the process, revolutionizing the technology of the world. Along the way the work has become a great deal more precise.

In the thirty-five years since the invention of the transistor, researchers have worked to make the device ever smaller, struggling to squeeze tens, hundreds, then thousands of transistors onto a single chip less than half an inch square. Now some scientists have succeeded in building chips with 256,000 devices, and the one-million mark may be less than five years away. Small has become much more than beautiful; it's become essential. As integrated circuits grow in complexity, the range of their uses broadens, and the race to develop new applications is on: from PAC-MAN to digital watches, from increasingly powerful radiotelescopes to household computers costing less than a vacuum cleaner. More and more often, microelectronics are making the big jump from laboratory bench to department-store shelf, and as research continues, scientists are gradually erasing the line separating science fiction from science fact.

At Yale this research stretches from individual devices to intricate combinations of circuits, from the near-atomic level to large-scale applications, and throughout, the challenges are as varied as the field. Within the laboratories and "clean rooms" of Becton Center, in the realm where the head of a pin looks like a football field and a human hair is as thick as a tree trunk, even the most straightforward experiment takes on new difficulty. Simply creating the components for these experiments poses a problem when they're on the order of one millionth of a meter in size, and developing new ways to combine devices when dealing with hundreds of thousands at a time can be insanely complicated.



The transistor is our friend. *At the foundation of today's integrated circuit technology is a collection of tiny and relatively simple building blocks called transistors. These devices are essentially switches. They enable a small current or, as in the case of the n-MOSFET transistor diagrammed above, a small voltage to control a much larger one, which in turn may be used to control other devices. When the transistor is off, voltage at the gate is zero, and although a large voltage is applied between source and drain, current cannot flow; both the silicon oxide (SiO_2) layer and the electron-poor p-type silicon are effective insulators. When the device is turned on, however, a positive charge applied to the gate attracts negatively charged electrons to the silicon below. These electrons are unable to penetrate the oxide layer. Instead they accumulate to form a "corridor" along which current can flow between n-type areas, completing the circuit and reducing the exterior voltage to zero.*

Right: Graduate student Eric R. Fossum checks the design of a microchip mask.

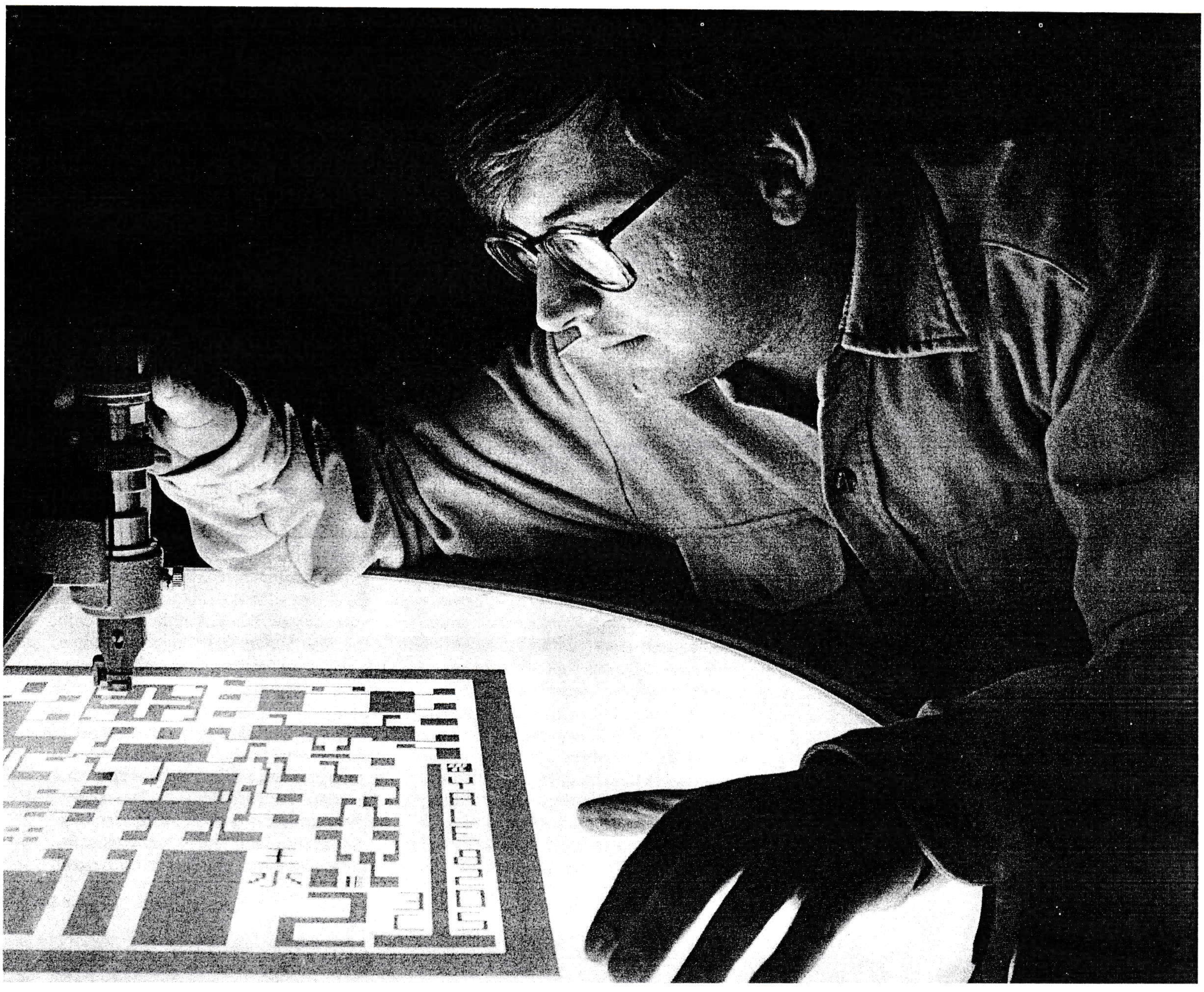
"It's a bit like building a cathedral with very small bricks," said Professor Werner P. Wolf, Chairman of the Council of Engineering. "The building blocks are very simple, but the whole edifice is impressive because you put them all together in a very clever way."

The foundation of this edifice is the element silicon, whose complex physical and electrical properties make the whole field of microelectronics possible.

Silicon is in a class of elements called semiconductors and, as such, sits somewhere between good electrical conductors like metals and good insulators. When treated with the appropriate substance, the so-called "dopant," it can be made extra rich in electrons or, conversely, in positively charged electron holes. These holes are not particles; rather, they are places where electrons are missing, but they behave and flow like positive particles. Thus, after "doping," the semiconductor can be made to carry a current. At the same time, if it is exposed to oxygen or steam at very high temperatures, it will form silicon oxide, an effective insulator. This combination of characteristics allows engineers to implant relatively simple patterns on a polished wafer of silicon, patterns of insulators, metal conductors, and doped areas, which function as transistors.

A transistor is basically a switch; it allows a small current to turn on and off a larger current. In principle it can be likened to a dam holding back a large reservoir of water. With relatively little effort, a workman could open a valve at the base of the dam and immediately release a torrent of water. Similarly, in the most common variety of transistor, a tiny initial current alters the electrical properties of the doped silicon, transforming it from nonconductor to conductor, and providing a ready pathway for a much larger externally supplied current to flow. One type of device—the metal-oxide-semiconductor field-effect transistor or MOSFET—is composed of two islands of negatively doped (n-type) silicon. The surface of the silicon wafer is covered with a layer of silicon oxide, and punching through this insulator into the two n-type islands there are two metal contacts, called the source and the drain. A third contact, called the gate, is laid on top of the insulator between the two islands, and although it is not in direct electrical contact with the silicon below, it can generate an electric field which controls the flow of current between the source and the drain.

To operate this transistor, a constant positive potential is maintained on the drain. This positive charge tends to attract negative electrons from the source, but because electrons cannot travel



through the electron-poor p-type silicon, no current flows. When a positive charge is applied to the gate, however, it induces an electric field in the p-type silicon just below it which pulls in electrons from throughout the p-type area. The electrons cannot penetrate the silicon oxide, so they accumulate at the interface of the semiconductor and its oxide, creating an electron "corridor" along which current can flow. Thus, when one turns on a voltage to the gate, a much stronger current flows from source to drain. When one switches off the gate voltage, the gathered electrons disperse, and the corridor disappears.

This sort of on-off switch, although not perhaps very exciting in itself, is the building block from which computers are made. Using a system of logic devised by British mathematician George Boole in the middle of the last century, problems can be expressed as a series of true-false questions. If one allows 0 to represent false and 1 to represent true, one can reduce mathematical calculations to base two and solve them through a network, albeit a very complex one, of on-off switches. Individual transistors can be combined into structures called Gates (not to be confused with MOSFET gates), which are designed to perform the logical operations "and," "or," "not and" (abbreviated as "nand") and "nor," accepting combinations of electrical impulses, 0's and 1's, and transmitting, in turn, a solution impulse, also either a 0 or a 1. Thus, an "and" Gate with, say, eight input leads would transmit a signal, a 1, only if all eight inputs were on; while an "or" Gate would transmit a 1 if any of its inputs were 1.

The trick in understanding the applications of this is to recognize the enormous number of Gates needed for complex calculations; each logical step requires its own combination of Gates; each mathematical or logical contingency must be broken down and dealt with by the proper array of transistors. Although Gates may be interconnected and used repeatedly in the course of a calculation, sophisticated computers require huge numbers of these devices. A single pocket calculator may have several different chips, each containing 16,000 Gates, while today's largest computers contain millions.

With numbers that high, it is clear why scientists are trying to squeeze more devices onto a single chip: the smaller they can make a computer, the faster it will be. Since a computer's speed is ultimately limited by the speed of light, the speed with which electricity can switch the transistors, the distance the current has to travel becomes a factor. Although to human beings the difference between a millionth of a second—the time it takes electricity to travel back and forth across a large room—and a billionth of a second—the time it takes to move a foot—is negligible, to a computer it offers the chance to make a few thousand more calculations.

T.P. Ma is helping to give computers this chance, working on problems related to the development of smaller devices, while at the same time offering his students the opportunity to try their own hands at making chips. Ma, an Associate Professor of Electrical Engineering, offers a course entitled Solid State Devices and Microelectronics to an assortment of undergraduates and graduate students each year. In the classroom and the lab the course lays out the techniques which—if one has steady hands, and one doesn't make any mistakes, and one is very, very careful—can produce the tiny, delicate hearts of computers.

"This course teaches students the fundamentals of integrated circuit design and technology, the principles of device physics, circuit fabrication," explained Ma. "Essentially the ultimate goal is to have them go through all the basic processing steps and then be able to characterize them from step to step: what has gone wrong and what has been done right."

The first project begins with a perfectly straightforward command: clean the wafer. It should be simple enough. The students begin. It isn't simple; it takes about ten separate steps.

"It's very important in microelectronics that you keep everything extremely clean," Ma observed. "The requirement here is much more stringent than what you might see in, say, a biology laboratory."

Working at the size of the smallest transistors, even a speck of dust is a potential enemy, so the finest work is done in a special "clean room," all participants carefully wrapped in pristine white smocks, caps, and shoe covers. In the laboratory the two-inch-wide circular wafer, sliced from a single crystal of silicon and polished to a dark mirror finish, is gradually shuttled from an acetone bath to water to sulfuric acid and back, then into a buffered oxide etching solution (BOE), composed chiefly of

T.P. Ma

hydrofluoric acid, which removes any residual oxide. After that it all begins again: water, acetone, and BOE. When the process is finished, and the chips are free of all surface impurities, the actual process of implanting the pattern begins. Because the details of the devices are too tiny to be etched by hand, the students must use photolithography to define the shapes of the devices.

First, the entire surface of the wafer is oxidized in a furnace at 1,000° C to produce a thin diffusion barrier which will be opened over the specific areas to be doped. A thin layer of a photosensitive chemical called photoresist is spread over the oxide to act as a type of photoemulsion. The intended design is inscribed on a glass mask, and when this mask is placed over the wafer and exposed with ultraviolet light, the wafer may be developed like a piece of film. The developer washes away all the unexposed photoresist, revealing the oxide beneath. When the wafer is then immersed in BOE solution, the exposed oxide is etched away. The remaining photoresist can be removed in a bath of acetone, leaving behind the desired pattern laid out in silicon oxide. If the wafer is then placed in a high-temperature furnace containing the dopant (boron or phosphorus, depending on the type desired) this material diffuses into the surface. The silicon oxide may then be dissolved in BOE solution, leaving behind the desired pattern of doped silicon.

Using a similar process of protecting, exposing, and developing, it is possible to lay down a finer layer of silicon dioxide, cut holes for electrical contacts, lay down metal leads for the source, drain, and gate, and make the necessary connections between devices. The completed wafer is then sliced up into individual chips and, under a microscope, the students cold-weld gold wire onto the metal contacts to connect the integrated circuit with the outside world.

In this class project there were about thirty devices on each chip; in large-scale integrated circuits (LSI) in industry there may be 30,000. Microlithography techniques suitable for the one don't suffice for the other. While constructing glass masks provides good training for students, industrial companies interested more in results than in training turn to computers for help. A circuit designer sitting at a computer terminal can instruct the machine to reproduce a device design any number of times, almost effortlessly laying out the thousands of individual transistors in the proper patterns. Once the computer has completed the design, it can create the appropriate mask, either by imprinting the pattern on a layer of photoresist using a precisely controlled flasher or by using an electron beam from a scanning electron microscope, which has a much smaller wavelength, to etch a design on a layer of electron resist. With photoresist and ultraviolet light, an engineer could produce clear images about one micron (a millionth of a meter) in size before optical diffraction blurred them. With an electron beam it will be possible to get a factor of 100 smaller.

As the devices get smaller, however, the thickness of the silicon-oxide insulator layer decreases proportionately until eventually, at a thickness of less than forty angstroms (four ten-billionths of a meter) its insulating properties break down. Electrons may be drawn right through the silicon oxide to the gate, nullifying the transistor's switching ability. But this "electron tunneling" effect has its own uses, which Professor of Electrical Engineering Richard Barker has been exploring.

"What we're looking for is a new device, one that does something new, and we're looking for new ways to study the interface between the oxide and the silicon," said Barker. "When we get to forty angstroms or less, the oxide is no longer a true barrier to the flow of current. We're trying to use that flow of current in device structures and study the characteristics of these oxides, especially near the silicon interface."

The new device that Barker and his team devised is an optical threshold detector, a switch that turns on whenever it detects a specified brightness of light. The traditional MOSFET optical detector, which has been in use for years, measures the number of photons of light that strike its surface and registers that as a charge which may be measured. What Barker's detector does, however, is remain off until the intensity of the light reaches a certain pre-set level, then turn on and stay on. Its effectiveness depends on Barker's ability to construct a layer of silicon oxide with great precision, since the threshold at which the switch operates is partly dependent on this thickness.

The threshold detector is composed of a layer of oxide less than forty angstroms thick on a silicon wafer, sandwiched between two electrical contacts. When a negative charge is placed on the oxide, it repels electrons and attracts positively charged holes to the silicon/oxide interface. Because the oxide layer is so thin, however, the holes don't stop at the interface; they migrate out through the layer while low-energy electrons migrate in. At the same time the tunneling holes are continuously replaced as normal thermal vibration in the silicon occasionally "pops loose" electrons, and the electric field carries the resulting holes to the oxide barrier. The system maintains a stable equilibrium.

When a light shines onto the wafer, however, this equilibrium collapses. The light energy excites electron-hole pairs in the silicon and begins to pop loose more and more positively and negatively charged particles, and as its intensity increases, the positively charged holes build up at the barrier faster than they can tunnel out. This accumulated positive charge alters the potential field across the wafer and produces a much higher potential difference across the silicon oxide than across the silicon. As a result, the low-energy electrons that have been tunneling in through the oxide begin to pick up more kinetic energy in their journey. When they hit the silicon layer, they frequently have sufficient energy to pop loose still more electrons and electron holes. As the positive charge build-up continues, the kinetic energy of the tunneling electrons rises, and they pop loose more and more electrons, which in turn create more positively charged holes. The result: positive feedback. An electron-tunneling current begins to flow through the silicon, and the switch is suddenly turned on.

Although the process by which these devices work may seem impossibly complicated, the potential applications are disarmingly simple. With large arrays of these tiny detectors, scientists will be able to produce instant contour maps which precisely divide a field according to brightness.

"If you wanted to look at some kind of image that had bright areas corresponding to water with sun reflecting on it, say, and dark areas corresponding to populated areas or forests, you could expose the image onto a plane of these devices," Barker explained. "You'd get thresholds everywhere. You'd produce something that would be all black in some areas and all white in others. Now you'd have a simple way to measure what fraction of the area was water."

In a more sophisticated usage these devices might also be adapted to provide a visual sense to robots and computers. Because these detectors essentially digitize visual images, converting shades of color into a series of black or white components, they can make the data more readily accessible to a computer's binary logic. Automated machinery would then be able to finely adjust its actions to react to its surroundings. But these applications are all just possibilities now.

"We don't know whether this will ever change the way the world lives or find its way into your neighborhood grocery store," Barker said. "But the fact is that it's a new device which we set out to invent. We knew what we wanted it to do, and we thought there was something in there that ought to be possible. It's only once in a lifetime that you set out to invent something to do a certain thing of this kind and it works. And what's really fun about it is that it involves quantum mechanical electron tunneling, and the highest level of capability of processing the semiconductor surfaces which we've had to develop."

In the realm of microelectronics, these techniques of processing and fab-

Richard Barker

rication are often as much of a challenge as the theoretical science behind a device. Surfaces must be uniform as well as thin, and all devices must be reproduceable. As the drive for smaller devices continues, these problems grow even more complicated. Associate Professor of Applied Physics Daniel Prober has spent much of his seven-and-a-half-year career at Yale exploring innovative techniques to get around the limitations of size inherent in photolithography in order to explore small-scale physics.

"In the early days we had fairly grungy equipment, and by and large we had to be very clever," Prober said. "Dick Barker gave me an old evaporator used for making metal films which he had gotten as an IBM hand-me-down. It was at least twenty years old then and would probably have been relegated to the junk pile by most sane people, but it was all I had and for five years we made very heavy use of it."

Prober wanted to build individual devices smaller than the 5,000-angstrom limit imposed by the wavelength of light, but all he had was a machine that was known to be incapable of making ultra-small structures. He considered the problem at length. Then one night, while lying in bed staring up at the play of shadows across his ceiling and walls, he found a solution. It was as simple as casting a shadow. He etched a step into the surface of a wafer, and because etching is not limited by diffraction, he was able to create a step 200 angstroms high. He evaporated a thin layer of gold onto the entire surface and then proceeded to etch away the unwanted metal using a high-energy ion beam. Instead of directing the beam straight down onto the surface, however, he angled it so that the step in the surface would cast a shadow, protecting the gold running along its base. In a process like scraping excess putty from a window, Prober was able to remove all the gold except a small, triangular wire no more than 200 angstroms wide.

"We bootstrapped ourselves down in size, which allowed us to conduct research over the last five years studying the limits of conductivity in very small systems," Prober observed. "It's been a very productive idea for making structures that are fairly simple, and there are a large number of interesting scientific experiments—in fact some of the crucial ones in solid-state science—that we now do with these techniques."

Although Prober's microfabrication technique was important in testing scientific predictions made about the behavior of metals, it was perhaps even more important as a demonstration of how technology and microfabrication techniques could be applied to pure science.

"In some ways it was regarded as the premier experiment demonstrating the applications of microfabrication technology to physics," he noted. "Until this time microfabrication technology was used to make devices, useful devices, profitable, small, sexy devices, but in fact it had not really been employed to do very much useful in science. The whole flow had been from science to technology. Our work was regarded as a major proof in principle that you could get something back for basic science."

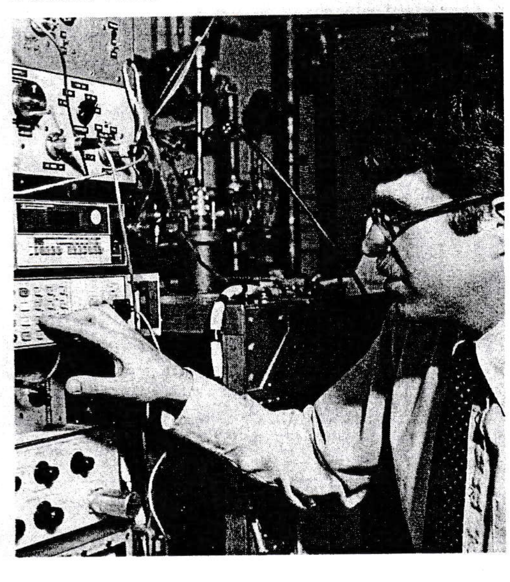
Part of what Prober got back was a technique for making superconducting tunnel junctions, tiny sandwiches of superconducting metals and insulating oxides which could be used to detect microwaves. These devices used electron-tunneling effects similar to those explored by Barker, but with different materials and under different conditions. At temperatures within a few degrees of absolute zero, the tin used in these tunnel junctions becomes a superconductor—a material with no electrical resistance. Two minute strips of tin are separated by a very thin insulating layer of tin oxide. Under normal conditions electrons cannot pass through this oxide, but when microwaves strike the tin strips, they generate an electron-tunneling current which can be measured and recorded. When incorporated into radio telescopes, these microwave detectors approach the absolute limits of accuracy and sensitivity set by the basic laws of physics.

Prober's work on microfabrication techniques, however, has applications beyond his own work in superconductivity. It dovetails closely with research that Professor Robert Wheeler, Chairman of the Section of Applied Physics, has been conducting for the past ten years. Wheeler has been studying the behavior of electrons, exploring how current is affected by temperature, magnetic field, and electrical field in MOSFET devices so small they are essentially one-dimensional. Using microlithography techniques he has created transistors that are about one-tenth the size made in industry, about 3,000 angstroms, and by conducting his experiments at extremely low temperatures, about the temperature of liquid helium, he has been able to minimize the effects of thermal and outside energy on the electrons running through the devices. Under these conditions, he found, a transistor does not operate simply as an on-off switch, but rather as a more complex switch with a series of on-off levels. Soon, with the addition of a recently purchased $85,000 scanning electron microscope, both Wheeler and Prober will be able to create devices another factor of ten smaller, and by using the electron beam from the microscope to create tiny uniform gates for his transistors, Wheeler can study the behavioral changes that occur in a single layer of electrons when the width of a device's electron corridor is decreased.

"There are other groups and research labs that are making small structures," said Prober. "But there's no other university lab in the country, with the possible exception of our friendly competitors at MIT and Cornell, that is now in a position to do the combination of studies that we can do. It's hot stuff. We're really on top of things here."

The work of Prober and Wheeler will ultimately result in smaller and smaller devices, but as the number of transistors squeezed onto a chip increases, the resulting circuits will offer their own challenges to the designers and engineers who attempt to apply them to concrete tasks. John Zornig, Associate Professor of Electrical Engineering, offers a glimpse of these difficulties to a class of undergraduate engineers every year in his "Digital Systems" laboratory. With nothing but a term of electronics under their belts, the students venture into the realms of system design, coding, and digital circuitry, and about a month into the course, they face their first project: the design and construction of an error-

Daniel Prober



detecting communication system. The finished products, which gradually take shape over several weeks of anguish and concentration, contain about thirty small, integrated circuits and two or three hundred wires. And more importantly they all work.

"These projects are orders of magnitude more complicated than anything they've done before," said Zornig. "One big part of the content of the lab is to teach them how to be orderly about design, not to do everything on little slips of paper or keep it all in their heads. It's too complicated. They're building things that are more complicated than they can keep in their heads at one time, and that is in fact what we do in electronics mostly: you build things that are more complicated than you can understand."

The second and final project is not much larger than the first, but it's a trifle more challenging. "Their second project lets the leash looser," Zornig observed. "In the first project I tell them what to design and they design it. In the second project I don't tell them what to design; I play customer. I say, 'There's a traffic intersection down at Temple and Grove. Make a light controller for that intersection that manages the traffic. The controller should adapt to the flow of traffic in some sensible way, adapt to rush hour, and so on. But other than that you have to think what the traffic controller should do and make it do that."

It's a course that faces young engineers with real problems and trains them by the traditional sink-or-swim method. Perhaps surprisingly, the overwhelming majority swim, and swim quite well. "Anyone who can get through that course is assured of a job," noted Engineering Council Chairman Wolf.

But while Zornig's laboratory course offers neophyte engineers their first exposure to complex design problems, it is still orders of magnitude simpler than the problems facing Zornig and the new group that the Engineering Department is assembling this year to expand the department's work into very large-scale integrated (VLSI) circuits. With the addition of Professor of Electrical Engineering and Computer Science Martin Morf last fall, the University has begun to lay the foundations for the study and development of circuits with hundreds of thousands of devices.

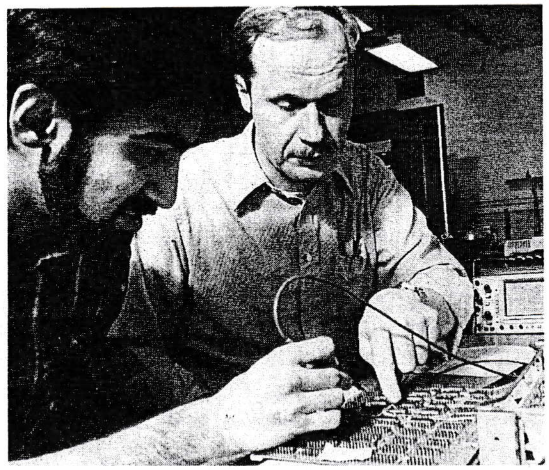
"The microelectronics program around here has for a long time been focused on single small devices," explained Zornig. "As we restructured the Engineering Department last year we tossed around a number of ideas for what direction to enlarge in. Ultimately we decided that a major direction would be very large-scale integration: the business of putting tens to hundreds of thousands of transistors on a single chip."

The field of VLSI breaks up into a series of smaller problems, all of which must be solved. There are materials problems—problems of ensuring that the very thin layers of oxides and metals are pure and uniform—and optics problems—making sure that you can plant all the needed information on a piece of silicon that small without letting any of the patterns go out of focus. But most of all there are design problems. If you're dealing with 100,000 separate devices that all have to be laid out on a single plane and that all have to be wired together without interfering with their neighbors, it means you had better be sure not to design yourself into a corner and discover at the last minute that somehow you have to connect two devices at opposite ends of the chip.

"As soon as you start building circuits that have to live in flat land on a little plane less than a millionth of an inch thick, then all of a sudden it becomes very important in your circuit design that you not crisscross all over this thing," Zornig noted. "So there is this second level to the design problem: how do you design a circuit such that you get all these hundreds of thousands of transistors to fit on the chip, and not have to have wires running from one end to the other? It's tricky."

Part of what the new group will be doing, and part of what Morf is already working on, is to attempt not only to build larger circuits, but to make them work faster.

"We are at most an order of magnitude away from the largest-scale integration we'll ever be able to achieve," noted Zornig. "We're within a factor of ten of the most transistors we can put on a chip, which is about a million, using current techniques. And if that's the best we can do, we have an idea of how fast we can get. If you then start asking, 'How can you get faster than that?,' well, you go back to the old World War II code breakers. A single human being could only crack codes at a set rate. How do you crack codes faster than that? You get a thousand guys and have them all try to crack codes. Well, you do the same thing with computers."



John Zornig, a student, and his project.

This is the notion of parallelism that Morf has been exploring. In the past, most problems have been designed to be solved sequentially, one step at a time, so that in a computer, the processors work on a problem one at a time. From a computer's point of view, where a microsecond is considered quite a lengthy pause, this method of operation can be time-consuming. The alternative is to devise an approach to the problem that would allow hundreds of processors to be working simultaneously.

"So far, the best that anyone has been able to do is to identify some problems that admit to this sort of structure and then build machines to do that," Zornig said. "Right now the state of the art of structuring problems is fairly ad hoc. We're in the process of tooling up, hiring, and setting up a graduate program, primarily to design algorithms and circuits to do problems in parallel. The program will be based on Morf's research, but it will have to go beyond that. By and large the thrust of the thing will be to find ways to build faster circuits. Our circuits may not move their bits around any faster, but the problem will get solved a lot sooner."

With the establishment of this group, the Yale Engineering Department will have people working on the whole spectrum of microelectronics, from the operations of the smallest devices to the creation of the largest circuits, from the intricacies of fabrication to the complexities of design. The goals are similar in all branches of research: smaller, faster, more complex—smaller devices, faster processors, circuitry capable of more and greater tasks. The pocket calculator with a brain the size of an aspirin tablet is just the beginning. Perhaps engineers in the future will find themselves wondering how many chips one can fit on the head of a pin. □