CHAPTER 3

Electronics Technology

Contents

	Page
Overview o...... ... . o.... e..... o.	**6 7**
Consumer Electronics	**68**
Semiconductors	**71**
Transistors	72
Integrated Circuits	72
System Design and the Microprocessor	73
Microprocessors and Memory	74
Learning Curves and Yields	76
Integrated Circuit Design.	77
Manufacturing Integrated Circuits	78
Future Developments	81
Computers •*,*,,.. ....***.. •.*.**... .	81
Types of Computers	82
Computers as Systems	83
Computer Software	85
Growth of Computing Power	87
Applications of Computers	90
Summary and Conclusions	92
Appendix 3A-Integrated Circuit Technology	93
Appendix 3B—Computer Memory	99
Appendix 3C--A Process Control Example	102

List of Tables

Table No.	*Page*
l. Examples of Solid-State Technologies Used in Information Processing	72
2. Cost Increases for Fine-Line Lithography	80
3. Characteristics of Different Categories of Digital Computers	83
4. Increase in Computing Power Over Time ....•.*..,,, •,,...,, .₀.	8a
5. Comparison of IBM 650 and Fairchild F-8 Microcomputer	88
6. Characteristics of Generations of Computer Technology	90
7. Typical Applications of Computing Capability	91
3A-l. Common Terminology and Classifications of Integrated Circuits	94
3A-2. Principal Types of Memory Circuits	96
3B-1. Computer Memory.	101

List of Figures

Figure No.	*Page*
1. Simplified Diagram of TV Receiver Componentry	69
2. Comparative Feature Sizes of Microelectronic Devices Such as Integrated Circuits	71
3. Controller for a Microwave Oven Based on Single-Chip Microcomputer	75
4. Increases in IC Integration Level	75
5. Schematic Learning Curve for the Production of an Integrated Circuit or Other Manufactured Item	76
6. Ranges in Cost and Time for Design and Development of Integrated Circuits	77
7. Steps undesigning and Manufacturing an Integrated Circuit	78
8. Step-and-Repeat Process of Photolithographic Pattern Formation on Silicon Wafer	79
9. Typical Microcomputer Intended for Personal and Small-Business Applications	84
10. Typical Minicomputer Installation Including a Pair of Terminals and a Printer	84
11. Data Processing Installation Built Around General-Purpose Mainframe Computer	84
12. Simplified Block Diagram of a Microprocessor System	84
13. Elements of a General-Purpose Digital Computer System	85
14. Relative Hardware and Software Costs Faced by Users of Larger Computer Systems	86
15. Minicomputer Price Trends	89
16. Parallel Decreases Illustrate How Costs of Computers Depend on Costs of Semiconductors	89
3A-1. Silicon Wafer Showing Patterns Formed by Photolithography	97
3A-2. Integrated Circuit Chip Before and After Final Encapsulation	98
3B-1. Performance of Computer Memory Alternatives	100
3C-1. Portion of Process Control System for a Chemical Plant	103

CHAPTER 3
Electronics Technology

Overview

This chapter outlines the technology on which the consumer electronics, semiconductor, and computer industries depend, covering them with enough depth to provide background for discussions later in the report concerning the role of technology as a force on competitiveness. Except for occasional examples, competitiveness itself is left to later chapters.

The primary function of electronic components and systems is to manipulate and transmit *information* in the form of electrical signals—either analog or digital, The transmission and utilization of electric power are integral to these processes, but constitute only secondary functions of electronic equipment. Even in the case of a 50,000-watt radio broadcasting station, the high power simply increases the area coverage of the information in the signal. The information manipulated and conveyed via an electronic system can range from a simple sequence of numbers—e.g., a zip code, or the balance in a checking account—to the sounds conveyed by radio, images such as television pictures or weather maps, or the information contained in radar or sonar signals.

Changes in electrical voltage are the most common carrier of information in electronic systems. In an analog system, the signal takes the form of a voltage or some other electrical parameter that varies continuously over a range, while *digital* information is encoded in the form of a string of binary "bits." Each binary bit can take on one of a pair of discrete values, again usually voltages. The magnitude of these is unimportant, so long as they can be distinguished from one another. A bit can be visualized as having values of "O" or "l, " or "+" as opposed to " – ." In a digital circuit, the signal normally takes the form of a string of discrete voltage levels—e.g., any voltage between –2 and i-l volts might represent a binary "O, " any value from + 2 to + *5* volts, a binary "1. "

Regardless of the simplicity or complexity of the information content in a signal, either analog or digital technology can, in general, be employed. The choice turns on the practical advantages and disadvantages for a given application. A complex system may use analog circuitry for some tasks, digital for others. In geophysical exploration, for instance, an analog signal—essentially a mechanical pressure pulse or sequence of pulses—is transmitted into a geological formation. The reflected pulse from the subsurface strata is sensed by transducers analogous to microphones. These transducers respond to the mechanical energy of the reflected pulse by generating a proportional analog electrical output. Analog-to-digital converters—typically integrated circuits (ICs)—then convert these signals to digital information that can be processed and analyzed by a digital computer.

Both analog and digital technologies have a place in the three sectors of the electronics industry covered in this report, But while most consumer electronics equipment is still based on analog technology—radio and TV receivers, phonograph records, magnetic tape players—virtually all computers process information in digital form. At the same time, computer peripherals such as terminals and printers contain analog circuitry, while digitally based consumer products are becoming more common. Semiconductor devices come in both analog (often termed "linear") and digital varieties. A few ICs combine analog and digital circuitry on the same "chip. "

Consumer Electronics

The most common consumer electronic products are radios, TVs, and audio equipment such as "stereo" systems. Electronic toys and games, electronic watches, pocket calculators, and home computers are other familiar examples. These are all "systems" in the sense that they contain more than a single electronic component, but some are much more complex than others. An electronic watch may consist of little beyond a single IC and a display—itself a solid-state device—plus a battery, Television receivers contain several hundred components. Video cassette recorders (VCRs) are complex mechanically as well as electronically,

Radio broadcasting provided the foundation for the development of the consumer electronics industry. Despite a real cost much higher than today, there were well over 10 million radio receivers in use in the United States by ***1930,*** with annual sales exceeding ***$1*** billion.[1]

Research and development (R&D) on television began in the 1920's, with limited broadcasting prior to World War 11. Large-scale commercialization had to await the end of the war, but by 1949 5 million TV sets were sold in the United States—all black-and-white. Color television—for which most of the early work was performed by RCA—followed the next year, but color TV sales in the United States did not pass the 5 million mark until 1967, and first exceeded black-and-white sales in 1972.[2]

With more than ***11*** million color sets sold in 1982, and about half as many black-and-white sets, the TV receiver remains the largest selling consumer electronics product, accounting for nearly half the dollar value of consumer electronics sales in the United States (ch. 4). Monochrome TV sales have been rather static for a number of years, with the market for color sets expanding only slowly; cable TV and direct satellite reception may spur future sales, but much of the growth in consumer electronics markets is now in new generations of products, notably VCRs. Still, in many respects —e.g., the relatively standardized design approaches and critical importance of production costs—color TV continues to typify consumer electronic technologies.

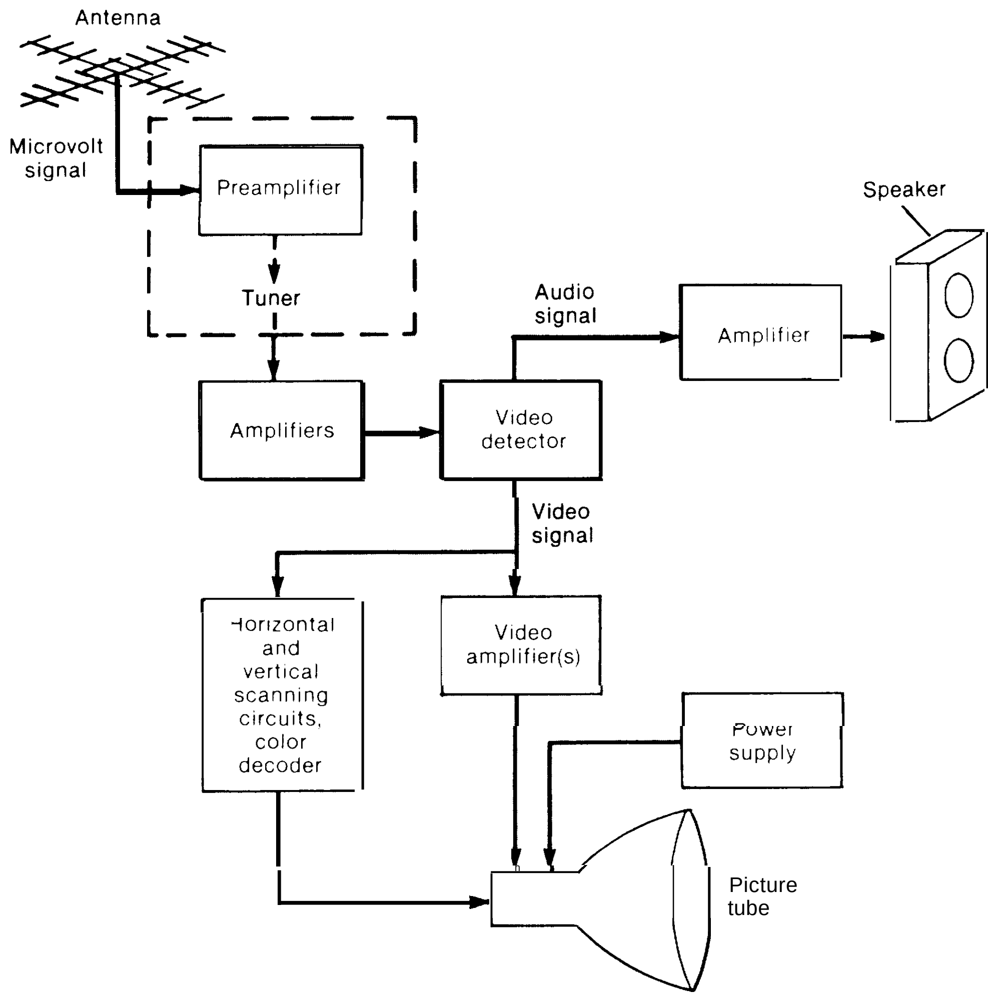
Television signals are broadcast via amplitude modulation of a high-frequency carrier signal, much like AM radio, But the bandwidth requirements for TV are far greater—about ***6*** MHz, versus 10 KHz for AM radio. Bandwidth, which is expressed in terms of frequency—6 MHz being equal to 6×10^6 cycles per second, 10 KHZ to 10×10^3 cycles per second—is a measure of the *rate* at which information can be conveyed, hence must increase with the *amount* of information in a signal. The pictorial image in a TV signal has a much higher information content than sound, hence television's high bandwidth requirements, In principle, the analog information in either a radio or a TV signal could be conveyed in digital form without changing the bandwidth requirements greatly.

A home antenna receives the amplitude-modulated TV signal at a microvolt level (1 microvolt equals 10^{-8} volts). To produce a visual image, this signal is amplified to control an electron beam which scans the front of the picture tube—or cathode-ray tube (CRT) —forming a new image **30** times each second (the number of frames per second can vary abroad). The circuitry in a TV receiver is quite complicated (fig. 1) and now entirely solid state (the chassis includes both discrete transistors and ICs) except for the picture tube, The CRT is the most expensive single component in the set, accounting for about 40 percent of the cost. Producing picture tubes is a highly specialized activity; smaller firms often buy CRTs from manufacturers such as Zenith or RCA,

Conventional picture tubes are not only bulky and expensive, but account for much of the power consumed by TV sets; in Japan and Europe particularly, consumer electronics manufacturers have devoted considerable effort to reducing power consumption. This not

[1] *Electronics* Apr. 17, 1980, pp. 44, 78.
[2] *The U.S. Consumer Electronics Industry* [Washington, L). C,: Department of Commerce, September 1975), p. 20,

Figure 1.—Simplified Diagram of TV Receiver Componentry



SOURCE" Adapted from K Henry (ed), Radio *Engineering Handbook,* 5th ed (New York McGraw Hill, 1959) ch 22

only reduces consumer electrical bills, but can lead to more reliable operation—one of the advantages that TVs imported from Japan have enjoyed over the years (ch. 6). The drawbacks of conventional picture tubes—principally bulk–have stimulated considerable R&D aimed at flat-screen television displays. Flat screens are not yet practical, but continued progress in solid-state technology will doubtless lead to eventual success.

Before 1960, most of the significant technical developments in television originated in the United States, but more recently important advances have come from Japan as well.[3] Television technology is now well diffused internationally, and no one country appears to enjoy a technological advantage. Product innovations continue to come from U.S. firms, but also from other parts of the world. European consumers, in particular, are often attracted by

[3] Ibid., p. 27; also "International Technological Competitiveness: Television Receivers and Semiconductors, Charles River Associates Inc., Boston, Mass., draft report under National Science Foundation grant No. PRA 78-20301, July 1979, app. 2A. Studies of technological developments in an industry such as this are inevitably judgmental, and often biased by familiarity with one's own country.

new and different product features such as multiple image displays (several channels shown simultaneously on one screen). These have become important to product development strategies of firms in Western Europe. Continued progress in large-screen projection TVs has been stimulated by competition between Japanese and American producers for what could be a large new market. Japanese firms have been leaders in reliability, and may have tended to emphasize R&D directed at automation—and at rationalization of the manufacturing process in general—more than European or American producers (ch. 6). In particular, Japanese TV makers were leaders in adopting transistorized chassis designs in the late 1960's and in automating the insertion of discrete components such as transistors, ICs, capacitors, and resistors into printed circuit boards.

Although TVs still account for much of the consumer electronics market, they are a mature product in the sense that most American homes already have one or more. Thus, the great proportion of sales are now supplements or replacements. The market for VCRs, in contrast, is expanding rapidly (ch. 4), US. sales of VCRs during 1980 were less than a million units—all imported; sales nearly doubled in 1981.

Video recording on magnetic tape was pioneered in the United States by Ampex.[4] Although Ampex and RCA continue to manufacture video tape recorders for broadcast applications, consumer VCRs were developed largely by Japanese firms (ch. 5)–which now build about 95 percent of the world's VCRs. In Europe, Philips has a few percent of the market, but all the VCRs sold in the United States come from Japan.

[4]"Interactions of Science and Technology in the Innovative Process: Some Case Studies, " final report, Battelle Columbus Laboratories, National Science Foundation Contract No. NSF-C 667, Mar. 19, 1973, ch. 1 2.

While the commercialization of VCR technology by Japanese manufacturers is one sign that Japan may be taking over product leadership in consumer electronics, video disks thus far present a mixed, perhaps contradictory, picture. The optical video disk system developed in Europe by Philips reached the consumer market first. In the Philips system, a laser reads the digitally encoded signal on a spinning disk; microscopic depressions in the disk represent binary "0s" or "1s." While an elegant technical achievement—and one with potential for high-density digital data storage of other types (e.g., in conjunction with computer systems)—the optical video disk sold in the United States by Magnavox has not been a commercial success. RCA's video disk, introduced early in 1981, functions on analog principles—more like a phonograph record. Yet a third system, developed in Japan by JVC, may eventually reach the marketplace. As compared to VCRs, disk systems are cheaper but can only play back, not record. While the technology is evidently in hand, it is too early to tell how large the market for home video disk players will be–e.g., whether it will rival that for VCRs—or which systems will survive in the marketplace.

A number of trends in consumer electronics —e.g., recent introductions of "component" TVs analogous to component stereo systems, along with games and low-end home computers that use a conventional television as the display–point toward the eventual development of more-or-less integrated home entertainment systems. Such systems might incorporate TV and audio reception and reproduction, including various kinds of information services, along with applications of computing capability—not only games, but recordkeeping, home security systems, control of household appliances, and regulation of heating, ventilating, and air-conditioning systems. Such developments do not depend heavily on technological advances except as low production cost is necessary for mass market acceptance.

Semiconductors

Strictly speaking, the term "semiconductor" refers only to the ***materials*** from which semiconductor devices are made. Such materials have electrical conductivities intermediate between good conductors like copper and insulators such as glass. Silicon is the most common semiconductor material—virtually all ICs, and most discrete transistors, are based on silicon, In this report, the term "semiconductor" will be loosely applied to the products of the "semiconductor industry" as well as to the materials that are the starting points for these products. The broader designations "microelectronics" or ''microelectronic devices' include semiconductor products-which have replaced vacuum tubes in nearly all applications—as well as other types of solid-state devices that process, manipulate, or display information.

The most familiar example of a vacuum tube application that solid-state technology has not yet been able to match is the CRT—not only the common TV picture tube, but the display screens of computer terminals. In other cases, solid-state devices are not only much smaller than vacuum tubes, but cheaper, more rugged, and longer lasting. They also offer higher operating speeds; indeed, on almost any measure of performance, microelectronic devices offer order-of-magnitude improvements over the components they have replaced. Modern digital computers would be quite impossible without semiconductors. Although a computer otherwise like current models could, in principle, function with tubes instead of ICs, such a machine would fill a building and probably not execute a single program without one or more tubes failing. Solid-state circuits have made ***practical*** many electronic systems that would earlier have been too big, too costly, or otherwise in fact unthinkable.

Although virtually all commercial microelectronic products are now made from semiconducting materials, considerable R&D has been devoted to classes of solid-state technologies with potential for transmitting and processing information based on principles other than conventional semiconductor physics. Such devices might function on magnetic or optical principles, rather than being strictly "electron- i c , although the materials involved are sometimes semiconductors. * Boundaries between electronic, magnetic, and optical technologies tend to blur as device technologies move toward microstructural and submicrostructural size ranges. (Microstructural sizes are large compared to interatomic distances but small compared to objects that can be easily seen and handled, like an IC chip itself; the feature sizes of microelectronic devices are currently in the range of 1 to 10 micrometers, or less than a tenth the diameter of a human hair—fig. 2.) Because solid-state devices based on magnetic or optical principles are often used as components in systems that are broadly electronic in nature, no fine distinctions will be made.

"Optical data transmission can give bandwidths much higher than electronic signals; this, along with the low raw material cost, is one of the advantages of optical fibers. Systems based on laser light sources, with optical fibers for signal transmission and thin-film integrated optical devices for signal processing, could replace many types of electronic circuits and systems.

Figure 2.—Comparative Feature Sizes of Microelectronic Devices Such as Integrated Circuits

	Representative feature sizes (design rules)	Examples	Dimension in micrometers
		Human hair	100
		Red blood cell	7
1978	5 micrometers		
1980	2 micrometers		
		Yeast cell	1
1985*	0.5 micrometer		
		Smallest bacteria	0.2
1990*	0.1 micrometer		
		Polio virus	0.01
			0.001
		Atom	0.0002

1 micrometer = 10^{-6} meter

*Projected.

SOURCE: Adapted from G. B. Larrabee, "Materials Characterization for VLSI," *VLSI Electronics, Microstructure Science*, vol. 2, N. G. Einspruch (ed.) (New York: Academic Press, 1981), p. 38.

Among these solid-state devices—table 1—are:

- transistors, ICs, light-emitting diodes (LEDs), all of which are semiconductors;
- bubble memories, which depend on the magnetic rather than the electronic properties of materials;
- liquid crystals (as in alphanumeric displays for watches or calculators), chemicals that change colors when their temperature changes;
- integrated optics, in which information is transmitted and processed in the form of light, Integrated optics and prospective future technologies such as organic semiconductors are not commercially important at present, but could have impacts on future success in international competition.

Transistors

While most of the fundamentals of semiconductor physics were known prior to World War II, the transistor itself was developed at Bell Laboratories after the war, and first demonstrated in late 1947.[5]

In contrast to passive electronic devices such as resistors, capacitors, and inductors—which can only respond to electrical signals—active circuit elements like transistors control and regulate the flow of electricity in a circuit. As a result, they can amplify electrical signals—a function that earlier could only be performed by vacuum tubes.

Transistors come in many varieties to serve different functions, just as for the vacuum tubes they superseded, To make a transistor, a semiconducting material such as germanium or silicon is "doped" with small amounts of elements—arsenic, boron, phosphorus—that locally affect its conductivity. The transistor, to the naked eye, is then just a small piece of, say, silicon with two or three wires attached. In fact, however, the purity, chemical composition, and perhaps crystal structure have been carefully tailored on a microscopic level.

Integrated Circuits

An IC is made by fabricating several circuit elements—transistors, capacitors, and such—on a single substrate. Integrated circuits were independently developed in the 1950's at Texas Instruments and Fairchild Camera and Instrument, still two of the leading semiconductor manufacturers in the United States (Fairchild is now French-owned). The two companies approached the problem quite differently during 1958-59, but their developments shared the common characteristic of an IC—two or more distinct transistors fabricated on a single substrate,[6] Thus they were monolithic circuits.

[5] W. Shockley, "The Path to the Conception of the Junction Transistor," *IEEE Transactions on Electron Devices, vol. ED-23,* 1976, p. 597; E. Braun and S. MacDonald, *Revolution in Miniature: The History and Impact of Semiconductor Devices* (Cambridge, Mass.: Cambridge [University Press, 1978), ch. 4.

[6] M. F. Wolff, "The Genesis of the I ntegrated Circuit, " *IEEE Spectrum,* August 1976, p. 45; Braun and MacDonald, op. cit., ch. 8. The depths of the transistors fabricated on a chip are small enough that ICs can be considered two-dimensional. Often the silicon substrate—a few millimeters on a side and less than a millimeter thick-is called a chip, as is the resulting circuit.

Table 1 .—Examples of Solid-State Technologies Used in Information Processing

Technology and examples	Description	Current status
Semiconductor electronics: Transistors Integrated circuits	Depends on electronic properties of semiconducting materials such as silicon, germanium, gallium arsenide.	Production
Magnetic devices: Bubble memories	Depends on magnetic rather than electronic properties of materials.	Bubble memories in limited production.
Solid-state optics: (sometimes called optoelectronics)	Depends on electro-optical properties of materials, some of which are semiconductors.	LEDs widely used for displays; integrated optics experimental.
Light-emitting diodes (LEDs)	LEDs are lighted when a current passes.	
Integrated optics	Thin-film devices in which signals are transmitted by light (photons) rather than electrons.	

SOURCE Office of Technology Assessment

Other types of ICs can also be built—e.g., hybrid or thin-film circuits—but monolithic devices comprise the bulk of production,

At present, the market for ICs is more than four times the size of that for discrete semiconductors. Because of this, and because very large-scale circuits pace the industry and are a major focus of international rivalry, this report gives much more attention to ICs than to other microelectronic devices.

Appendix 3A discusses IC technology in some detail. The significance of the technology itself for international competitiveness resides largely in the commercial advantages that can accrue from innovative and/or widely accepted chip designs (ch. 5), as well as from mastery of processing technology. Being first on the market with a new design gives a firm the opportunity to build market share before competitors can offer similar products. The advantage of a particularly well-accepted design is that it may become a de facto industry standard. A manufacturer whose design becomes such a standard not only has the assurance of a relatively large and stable market, but also the prospect of a broad range of licensing and/or second-sourcing agreements. The firm may also get a headstart in the competition to design the next generation replacement. Processing capability is just as important to competitive success as design, because advanced circuit designs are often limited by what can be built at acceptable yields. (The yield is the fraction of "good" circuits coming off the production line.) In semiconductors, advances in circuit design and in processing capability are interdependent to a greater extent than in almost any other industry.

The first of the two major types of ICs to be developed, bipolar, has been replaced for many applications by MOS (metal oxide semiconductor, app. 3A). Over the course of the 1970's, MOS technology—which is denser, cheaper, and consumes less power, but which does not offer speeds as high as bipolar—became dominant for large-scale integration (LSI). Very large-scale integration (VLSI) will be mostly MOS. Firms that were slow to master MOS tended to fare poorly in sales growth and profitability over the past decade. For the foreseeable future, competition in ICs will continue to center around MOS devices.

System Design and the Microprocessor

In designing a digital system, the engineer has several options:

1. to assemble a number of ***standard logic*** circuits like those of the transistor-transistor logic family described in appendix 3A;
2. in cases where large production volumes are anticipated or performance requirements are specialized and demanding, to call for one or more custom ICs;
3. to use a standard microprocessor or microcomputer with software written for the particular application.

Assemblies of standard logic circuits—typically small- or medium-scale ICs—may be economical in limited production volumes, despite relatively high design and development costs. In such cases, the system is implemented in *hardware---e.,.,* its functioning can only be altered by changing the circuit components and/or their interconnections.

Specially designed custom circuits—analog as well as digital, bipolar as well as MOS—have a place in high-volume applications ranging from consumer products like TVs and electronic watches to telecommunications systems. They are also employed where performance requirements such as operating speed cannot be met in other ways-e. g., in some military and aerospace applications, or in mainframe computers. Custom circuit design is expensive: hundreds of thousands of dollars, sometimes running into millions. Here the designer is also working in hardware, but new custom hardware rather than standard ICs.

In contrast, when a system based on a ***micro-*** processor or microcomputer is designed, the logic is implemented largely through *software* —i.e., a computer program stored in memory. Although ***microprocessors*** and single-chip

99-111 0 - 83 - 6

microcomputers can function as central processing units for general-purpose computer systems—such as the small machines sold by Apple or Radio Shack—they were originally conceived as replacements for custom ICs to circumvent the high costs of designing, developing, and producing custom parts, The first commercial microprocessor was introduced by Intel Corp. in late 1971 to implement the arithmetic functions in an inexpensive calculator. Faced with the request of their Japanese customer for a group of custom chips to be used in a line of calculators, Intel instead proposed a simple 4-bit microprocessor chip.[7] Rather than hard-wiring the operations required for the different calculator models—addition, multiplication, printing, and so on—software programs permanently stored in memory implemented these functions. Money was saved in design and production compared to the custom IC alternative.

Subsequent experience has shown that microprocessors may prove the low cost alternative for systems that could be built with as few as two or three dozen standard logic circuits. Figure 3 illustrates a typical application of a microcomputer, control of a microwave oven, where the chip contains memory for program storage along with a processor. In such an application, production volumes might be high enough to justify a custom LSI chip design—tens of thousands of identical parts, at a minimum, are normally called for. But the microprocessor/microcomputer alternative has a big advantage in flexibility; design changes are simple, different models simply need different programs. And, in the microwave oven example, there is no need for high performance.

[7]R. N. Noyce and M. E. Hoff, Jr., "A History of Microprocessor Development at Intel Corporation, " *IEEE MICRO,* February 1981, p. 8. Parallel developments took place at Texas Instruments,

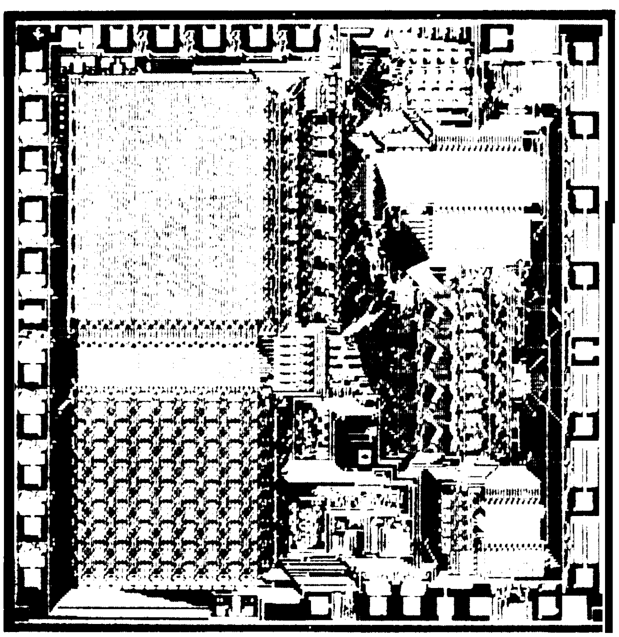


Photo credit Texas Instruments

A single-chip microcomputer

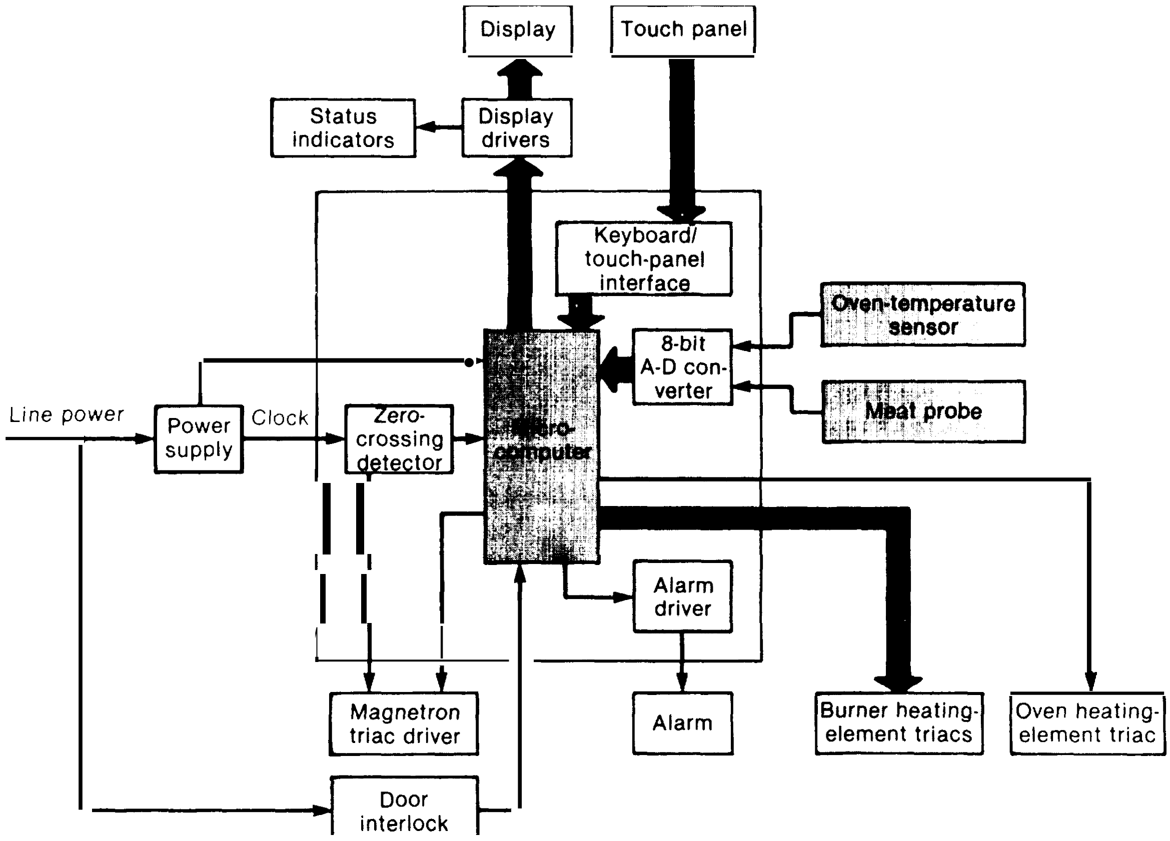
Before the advent of the microprocessor, system designers had only two choices: assemblies of standard parts or custom ICs. The microprocessor/microcomputer introduced a third option, one that proved highly attractive. As a result, several hundred different models of microprocessors and single-chip microcomputers are now marketed (many differ only in details), and custom microprocessors are sometimes designed for special applications.

Microprocessors and Memory

Microprocessors cannot be used by themselves; they must be supported by other chips, at a minimum for program storage. Memory circuits are described in some detail in appendix 3A, particularly table 3A-2. Memories, in fact, comprise the largest single market category for ICs; the majority go into general-purpose computer systems, but large numbers are also used in dedicated applications of microprocessors and microcomputers (i.e., applications where the computing function is invisible to the user of the system).

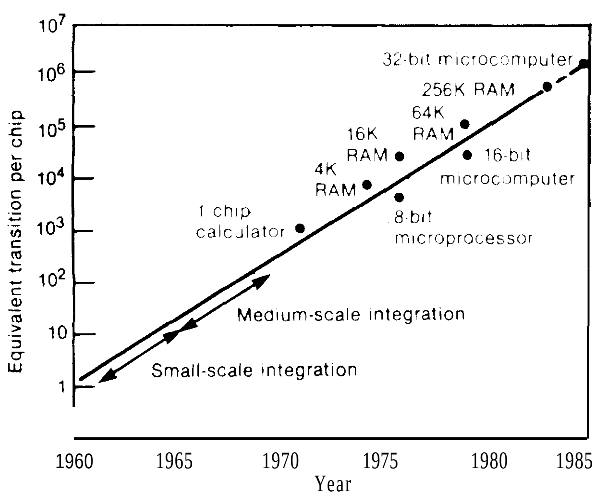
Technological progress is easier to measure for memory circuits—e.g., RAM chips (random access memory)—than any other type of IC. Densities have increased steadily over time—figure 4–while the cost of a chip has remained roughly the same, As a result, the cost per bit of information stored goes down, This opens up new applications, not only for memory circuits, but for IC-based systems of any type that

Figure 3.—Controller for a Microwave Oven Based on Single-Chip Microcomputer



SOURCE R Walker 'Analog/Digital LSI ' 1979 Electro Professional Program, New York, Apr 2426, 1979

Figure 4.— Increases in IC Integration Level



SOURCE D Queyssac, "Projecting VLSI's Impact on Microprocessors," IEEE Spectrum May 1979 p 38

rely on memory—notably microprocessor systems. Likewise, as semiconductor memory becomes cheaper it will continue to substitute for alternative storage media such as magnetic disks in general-purpose computer systems. It was widely noted in the early 1970's that the cost per bit of memory had fallen below the cost of a jelly bean. According to some estimates, a jelly bean (at 1¢) may buy as many as 1,000 (1K) bits of memory by 1990.

Memory and microprocessors are the most visible products in domestic and international competition. While it would be wrong to consider these the only important categories, they do constitute half the total market for ICs. Moreover, significant advances in both device technologies and process technologies have often found their way into production via circuits of these types.

Learning Curves and Yields

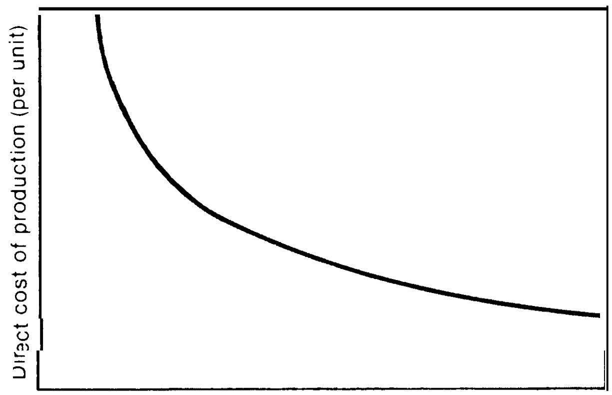
Making ICs is demanding, more so at higher levels of integration, Forty or more processing steps may be required for a VLSI chip, a figure that will continue to grow. Designing VLSI circuits is also complex, and becoming steadily more time-consuming and expensive, while product design and process design go hand-in-hand. For consumer electronic products like TVs, decisions on product features often hinge on the *costs* of production; for a new IC, the first question is: Can it be made at all?

Once a semiconductor firm has designed an IC and carried it through the pilot production stage they can normally assume that production costs, even if initially high because of low yields, will decrease over time. Figure 5 is a schematic learning *curve,* sometimes called an experience curve, showing cost declines with cumulative production volume. Learning curves typical of IC manufacture show that when cumulative production doubles, costs decrease by about 28 percent.[8]

Learning curves as in figure 5 apply to manufactured products of many kinds, but their impacts on pricing decisions have been particularly noticeable among semiconductor firms; they are a major factor in forward-pricing—setting prices below the initial costs of production to gain market share, Because firms feel confident that costs will decrease as production experience accumulates, forward-pricing has been a common competitive tactic,

These cost declines—which can be considered equivalent to increases in productivity—stem from much more than simple learning or experience by the labor force; other causes include better equipment performance and utilization, greater understanding and control of

Figure 5.—Schematic Learning Curve for the Production of an Integrated Circuit or Other Manufactured Item



SOURCE Off Ice of Technology Assessment

the many steps in the production process, smoothing of work flows, and perhaps changes in the design of the part itself. Control of the process is particularly important, and often depends on subtle variations in parameters influencing phenomena such as diffusion, etching, or polymerization (of photoresists—IC fabrication steps are described in more detail below). In many cases, the physics and chemistry of such phenomena are only poorly understood, and cannot be modeled theoretically; process control models tend to rely on empiricism, hence on accumulated experience, Denser and more complex ICs call for a better grasp of processing fundamentals.

In general, learning improvements show up as increased yield, the percentage of chips that pass final test and function satisfactorily. When a new IC goes into production, the yield is generally low—perhaps only a fraction of a percent—but rises as experience accumulates and processing can be better controlled. As a rule-of-thumb, products are seldom marketed, except for sampling purposes, until yields have risen to about 10 percent—which may take as much as a year of production-line experience.[9] For mature products, yields can rise to well over 50 percent, Increased yields are a power-

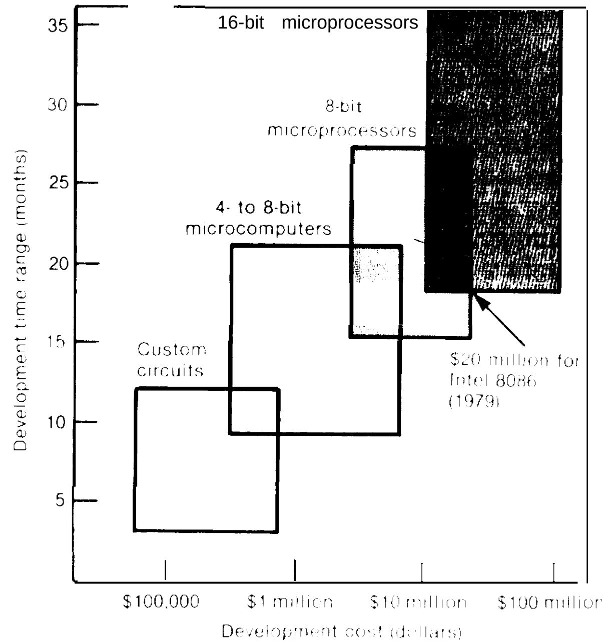
[8] "Boom Times Again for Semiconductors)" *Business Week*, Apr. 20, 1974, p. 65; *A Report on the U.S. Semiconductor Industry* (Washington, D. C.: Department of Commerce, September 1979), pp. 48-50. The 28 percent figure is an average from which the cost experience for a given IC can deviate substantially. Production volumes typically rise rather slowly at first, because it takes time for customers to design the new part into their systems. In comparison to other types of manufactured products, learning curves for semiconductors are not particularly steep, but continue to fall over very long production runs.

[9] R. Bernhard, "Rethinking the 256-kb RAM, " *IEEE Spectrum*, May 1982, p. 46.

ful force in driving down the costs and prices of ICs; in effect, doubling the yield halves the production cost.

Because processing capability is so critical to commercial success, the specialized equipment used in fabricating ICs is one of the keys to competitive ability. Much of this equipment is designed and built by independent suppliers —many of them American firms—selling to customers throughout the world (ch. 4). The technological capability and competitiveness of the portion of the U.S. electronics industry that designs and manufactures equipment has been just as important to the international position of the United States in semiconductors as the efforts of semiconductor firms themselves. Because of the interrelations of device technologies and process capability, and the dependence of costs and yields on process control, a number of the more important steps in producing ICs—beginning with circuit design—are described in more detail below.

Figure 6.— Ranges in Cost and Time for Design and Development of Integrated Circuits



SOURCE: P. M. Russo, "VLSI Impact on Microprocessor Evolution, Usage, and System Design," *IEEE Transactions on Electron Devices*, vol. ED-27, 1980, p. 1339.

Integrated Circuit Design

The task of the circuit designer is to define an arrangement of circuit elements—transistors, capacitors, logic gates, interconnections—that will satisfy the functional requirements of the IC. The more complex the circuit and the higher the level of integration (a 64K RAM contains more than 100,000 circuit elements) the more difficult the designer's job, and the higher the cost of design and development. Figure 6 illustrates ranges of development time and cost —including hardware, software, and peripheral chips—for several types of ICs.

As a rule-of-thumb, circuit design costs historically have been rather stable at about $100 per gate. Thus, a microprocessor with 10,000 gates will have a hardware design cost of perhaps $1 million, and may represent 10 man-years of effort. Software costs add to this. It may well be possible to make chips with 1 million gates within a few years, but the costs of designing them will be prohibitive unless the design costs on a per-gate basis can be reduced; one estimate has been that an IC with a density of a million devices would require about ***200*** man-years for design and development using the conventional methods of the past decade.[10] Computer-aided design—i.e., the use of specialized computer programs by the design engineers—is the principal hope for cost reduction, and is increasingly necessary just to handle the logical complexity as the number of devices per chip goes up. Designing a microprocessor with ***100,000*** transistors would ***be*** impractical without computer aids. R&D aimed at more regular —even modular—chip designs is also underway, again intended to reduce the time, hence the cost, of IC design. Modular approaches are particularly attractive for custom logic circuits.

As pointed out above, the microprocessor itself originated as a way to reduce the costs of custom circuit design; in essence. choosing a microprocessor means replacing hardware design by software design, and in many cases lowers costs, But as logic complexity goes up,

[10] C. L. Hogan, cited by F. Ogden, "Audience Gives Mixed Reception," *Electronics Weekly*, Mar. 28, 1979, p. 5.

the software costs for programing the microprocessor escalate rapidly. In part, this simply reflects the more sophisticated processors—e.g., a 16-bit rather than a 4- or 8-bit device—needed for demanding applications. Although software production can also be automated, the two basic paths toward implementing logic—hardware via custom chip design, or software via a standard microprocessor with a program embodying the logic—will continue to compete. Design cost, flexibility, and performance are all factors. But if computerized design aids for hardware and software advance sufficiently far and in tandem it may eventually make little difference, perhaps even to the designer, whether the logic is embodied in hardware or soft ware.

Although much of the engineer's work revolves around the logic that the circuit will implement, IC design also calls for intimate knowledge of processing and fabrication[11]—figure 7. Designs that can be implemented in n-MOS might be impossible in c-MOS (see app. 3A, table 3A-1, for an explanation of the types of MOS devices). One company might have n-MOS process capabilities beyond the reach of another. The design team must consider factors such as the spacing between transistors and the widths of the lines that interconnect them. In contrast, when designing a system to be built from discrete components it was often enough to be familiar with the performance characteristics of off-the-shelf devices. At the same time, IC designers—typically electrical engineers—must be at home with the logical concepts and software orientation of the computer scientist; the need for software skills will grow as ICs come more and more to resemble integrated systems. This melding of hardware (including process technology) and software skills, and the rapidity of technical change, make unusual demands on the people who fill such jobs—one reason that the electronics industry has been experiencing shortages of qualified engineers (ch. 8). Although neither the circuit designer nor the process specialist can be fully conversant with *all* aspects of IC design and manufacture (fig. 7), some knowledge of each is needed.

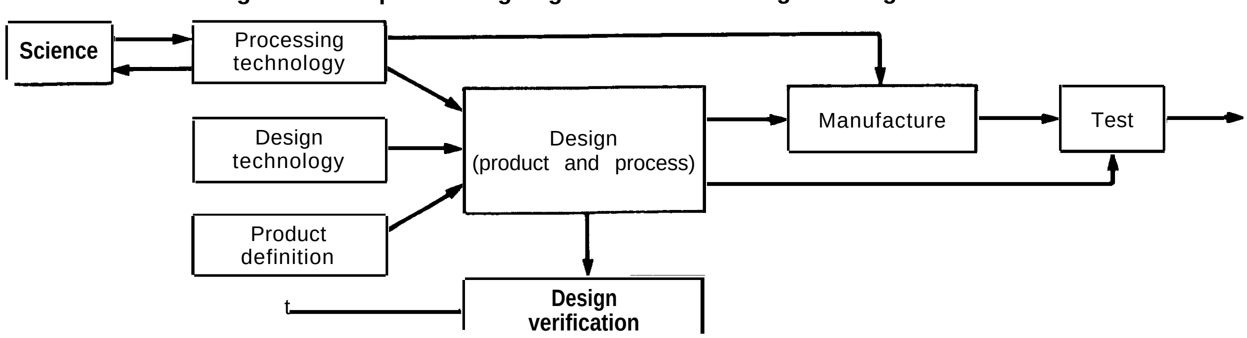
Manufacturing Integrated Circuits[12]

The design process culminates in a pattern or layout-a large drawing, several hundred times the size of the circuit itself—that must be translated into the "tooling" for producing the chip. In simple terms, the procedure for making an IC resembles a series of photographic processes—lithographic patterns are created in layers on a silicon wafer. Each layer is made by exposing a polymeric chemical called a photoresist to light or other radiation, the light

[11] P. W. J. verhofstadt, "Evaluation of Technology Options for LSI Processing Elements, " *Proceedings of the IEEE,* vol. 64, 1976, p. 842; C. Mead and L. Conway, *Introduction to VLSI Systems* (Reading, Mass.: Addison-Wesley, 1980).

[12] See, in general, F. W. Voltmer, "Manufacturing Process Technology for MOS VLSI, " *VLSI Electronics: Microstructure Science, vol.* 1, N. G. Einspruch (cd.) (New York: Academic Press, 1981), p. 1.

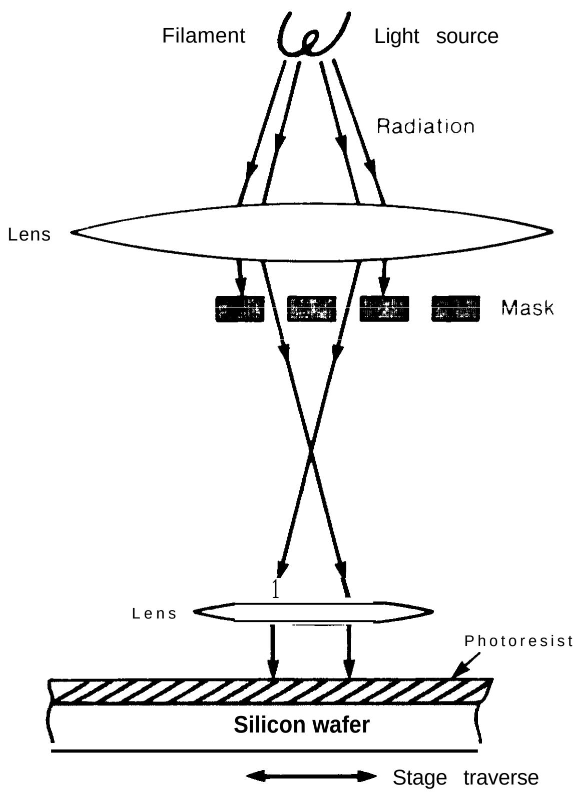
Figure 7.—Steps in Designing and Manufacturing an Integrated Circuit



SOURCE' G. Moore, "VLSI: Some Fundamental Challenges," *IEEE* Spectrum, April 1979, p.30.

passing through a grid-like mask as shown schematically in figure 8. More than a dozen such masking steps may be needed to build up a VLSI part. Many other processes besides lithography are involved in IC fabrication, with more detail given in appendix 3A, but lithography is critical for future increases in circuit density. While advances at many stages in the manufacturing process take place in an interdependent way—e.g., laser annealing is replacing furnace annealing because it does a better job of restoring the crystal structure of the silicon which is disturbed by ion implantation —lithography is the major factor in determining how small individual devices and interconnections can be on a production as opposed to laboratory basis. Already, transistors can be packed much more closely in an IC than neurons are packed in the human brain; it is the technology of lithographic processing that makes this possible.

Figure 8.—Step-and-Repeat Process of Photolithographic Pattern Formation on Silicon Wafer



SOURCE Office of Technology Assessment

Lithographic line widths in production ICs have been reduced an order of magnitude over the past decade, from about 20 micrometers in the early 1970's to 2 to 4 micrometers currently. (A micrometer is about 40 millionths of an inch,) Thinner lines give higher operating speeds as well as denser packing. The 16K RAMs designed in the early to mid-1970's were based on 5 micrometer "design rules," 64K RAMs on 3 micrometer rules (design rules, which are directly related to lithographic line widths, comprise the full set of geometric constraints that designers follow). The next-generation 256K RAMs are based on 1.5 to 2 micrometer design rules.[13] Continued progress in reducing line widths—more generally, feature size—is thus a major driving force in moving further into VLSI. For this reason, a principal R&D target of the Defense Department's Very High-Speed Integrated Circuit (VHSIC) program has been lithographic technology—an objective paralleling commercial R&D efforts, one reason the program is likely to have a positive effect on nonmilitary portions of the semiconductor industry. The VHSIC program goals include two stages of lithographic improvements: the first stage calling for line widths of 1,25 micrometers; the second, lines of 1 micrometer and below.

While feature sizes of ½ to 1 micrometer are well above the range for which the physics of electron devices will begin to constrain performance, such feature sizes do demand significant developments in lithographic capability, particularly for mass production .[14] In the past,

[13]"Rethinking the 256-kb RAM, " op. cit. On design rules, see Mead and Conway, op. cit., p. 47.

[14]I. E. Sutherland, C. A. Mead, and T. E, Everhart, *Basic Limitations in Microcircuit Fabrication Technology*, Defense Advanced Research Projects Agency Report R-1956-ARPA, November 1976; R. W. Keyes, "The Evolution of Digital Electronics Towards VLSI," *IEEE Transactions on Electron* Devices, vol. ED-26, 1979, p. 271, The ultimate limit to reductions in the sizes of electron devices will perhaps be thermal noise, although a variety of practical concerns may intrude first. Feature sizes are likely to decrease to the range of 0.01 to 0.1 micrometer before fundamental physical limitations are encountered.

visible light—generally ultraviolet—has been used to expose photoresists (fig. 8). But even deep ultraviolet, which has a wavelength of about ½ micrometer, cannot produce line widths much below a micrometer because optical considerations limit the lines to about twice the wavelength of the radiation.

To achieve 1 micrometer lines with visible light requires very sophisticated lithographic equipment—positioning and layer-to-layer registration of the sequential masking steps must be held to a small fraction of a micrometer. A single machine of the direct-step-on-wafer type diagramed in figure 8 now costs about half a million dollars—table 2. As the table shows, the earlier generations of equipment replaced by direct-step-on-wafer machines were much less expensive. Finer patterns will be still more costly—whether the technology of choice is electron-beam lithography (table 2), X-rays, or ion beams (see app. 3A). The rapidly rising costs of IC processing equipment—whether for lithography, for ion implantation, or for testing—along with higher costs for design and development, are the most important causes of the rapidly increasing capital intensity in the semiconductor industry (ch. 7). Entry costs are now roughly $50 million, versus $5 million to $10 million in the early 1970's.

Some firms have already moved to electron-beam lithography for critical circuit layers. Electron-beam lithography is also a routine tool for making masks. X-rays and electrons have wavelengths much less than light, and so offer greater resolution—at the expense of high first cost for the equipment, and low production rates.[15] Because of its importance in driving IC technology, R&D on high-resolution lithographic techniques (see app. 3A) has been a principal target of government-funded programs in other countries—e.g., Japan's VLSI project—as well as the VHSIC program funded by the U.S. military.

[15] G. R. Brewer, "High Resolution Lithography," *Electron-Beam Technology in Microelectronic Fabrication*, G. R. Brewer (cd.) (New York: Academic Press, 1980), p. 1.

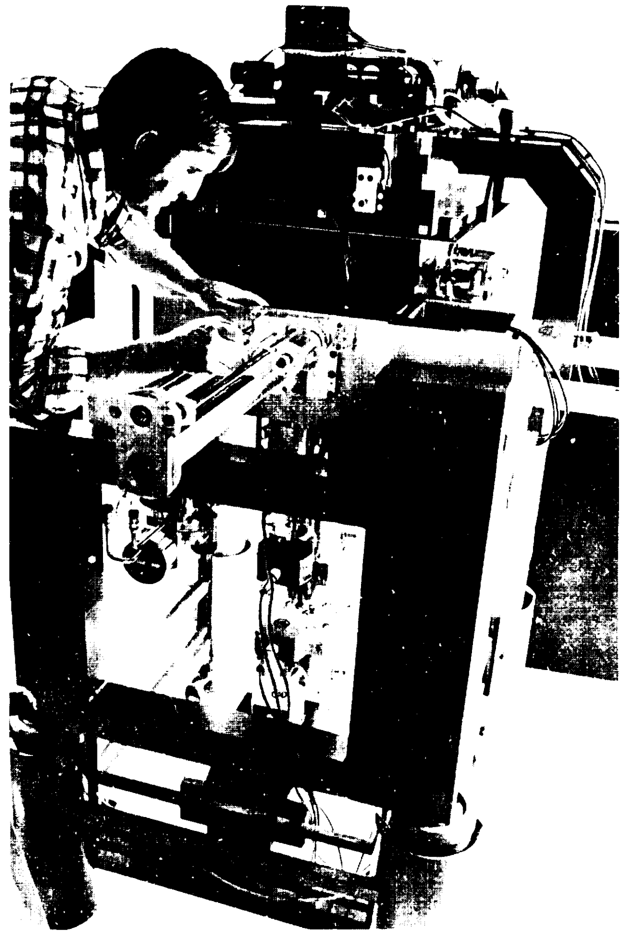


Photo credit GCA Corp

Electron-beam lithography system

Table 2.—Cost Increases for Fine-Line Lithography

Lithographic system	Line width (micrometers)	Throughput (wafers per hour)	Approximate cost per system	Approximate capital requirements for production capacity of 1,000 wafer starts per week
Light				
Contact printing	10	60	$15,000	$30,000
Projection	2-5	60	$240,000	$400,000
Direct-step-on-wafer.	1-2	30	$480,000	$1.6 million
Electron-beam	0.5-1.0	6	$1.5 million	$25 million

SOURCE Adapted from A J Stein, J Marley, and R Mallon, "The Ire-pact of VLSI on the Automobile of Tomorrow," *VLSI Electronics Microstructure Science*, VOI 2, N G Einspruch (ed) (New York Academic Press, 1981), p 295

More cost comes with the "clean rooms" needed for VLSI processing. Even micrometer-size dust particles can ruin the lithographic patterns; cleanliness is vital to high yields. In a clean room, the air is filtered and people must wear special clothing. As circuits become denser, and feature sizes smaller, not only is cleanliness more important, but the whole range of processing equipment used in making ICs becomes more sophisticated and expensive, adding to the capital requirements for semiconductor manufacturing, a matter discussed in chapter 7.

Future Developments

Semiconductor devices need not be based on silicon. One alternative is gallium arsenide, a material that offers considerable potential for improvements in packing density and speed—one to two orders of magnitude compared to silicon—but is still largely a laboratory technology. Whether gallium arsenide circuits will become commercially important depends on rates of improvement compared to silicon, and also on developments in other prospective technologies-e. g., Josephson junctions. Josephson device--also experimental, and much further from demonstrated practicality than gallium arsenide ICs—promise still better speed and density.

To illustrate the importance of speed, as well as power consumption, consider the technology embodied in a current-generation mainframe computer. The central processor for one such computer—the Amdahl 470-V—employs 1,680 ICs, with a total of about 100,000 logic gates. As is typical in large computers, the chips use silicon bipolar technology to give high computing speeds. Replacing these bipolar chips with gallium arsenide may offer the potential for increasing computational speeds by a factor of 10 to 100, and reducing the power consumed by the processor from about 3,000 watts to perhaps 30 watts—less than most light bulbs.[16] Comparable improvements in other applications of ICs carry implications for competitive trends in many industries if some companies or some countries manage a headstart in reducing such technologies to practice.

[16] R. C. Eden, B. M. Welch, R. Zucca, and S. I. Long, "The Prospects for Ultrahigh-Speed VLSI GaAs Digital Logic," *IEEE Transactions on Electron Devices*, vol. ED-26, 1979, p. 299 About 1 percent of the electricity consumed in the United States now goes to computers, mostly for cooling—"CBEMA Prediction: Say Energy Crunch Could Cut EDP Growth Rate 50%," *Electronic News*, Mar. 17, 1980, p. 32. On Josephson junctions, see J. Matisoo, "Overview of Josephson Technology Logic and Memory," *IBM Journal of Research and Development*, vol. 24, 1980, p. 113.

Computers

Computer technology has many roots, including military needs during the Second World War for fire control tables and other complex and/or repetitive computations. The United States had no great advantage over other nations during the early development of computers; significant innovations also originated in several European countries, particularly Great Britain.[17] But as computing technology progressed, the lead swung decisively to American firms, much as happened over roughly the same period of time for semiconductors.

The Bureau of the Census was an early nonmilitary customer for American computers, census data processing requirements remaining a typical example of computer applications. When a Univac I was delivered to the Bureau of the Census in 1951, some observers predicted that the market for digital computers might eventually total a dozen; a few years later, when sales to private industry began, the estimates were that the potential market in the

[17] *Gaps in Technology: Electronic Computers* (Paris: Organization for Economic Cooperation and Development, 1969), p. 61. For a concise summary of developments during the first three generations of computing technology, primarily in the United States, see S. Rosen, "Electronic Computers: A Historical Survey," *Computing Surveys*, vol. 1, 1969, p. 1.

United States consisted of perhaps 50 corporations.[18]

Needless to say, as computer technology advanced many more firms became customers, and the ranks of computer manufacturers swelled. Among the entrants were a number of companies that had become established in the office equipment market. International Business Machines Corp., Burroughs, and National Cash Register (now NCR) joined firms like Univac that had been set up specifically to manufacture digital computers. By the end of the 1960's, computer applications had spread well beyond numerical computations and data processing. The great part of computing power is still devoted to data processing for business and government—accounting, sales, production, inventories, recordkeeping of all kinds—and to scientific and engineering calculations. In addition, many individual computers, mostly microprocessors and microcomputers, now perform "invisible" functions in applications ranging from aircraft flight control systems to the microwave oven example shown in figure 3.

The spread of computing power has sometimes been technology-driven, sometimes driven by user demands. Technology-driven developments arose when more computing capability was available than people knew how to use productively—i.e., before the applications were well-defined. Under these circumstances, the availability of more powerful machines or greater performance per dollar tends to generate new applications, or, more broadly, serve needs earlier unmet. As in many instances of technological change, what the technology could do, at any given time and for a given cost, evolved in conjunction with applications, with one or the other temporarily in the lead, Much the same has been true for ICs. In the period when demand from military and space programs in the United States was high, the market drove the technology; but leaders in the semiconductor industry have periodically worried that applications for the full capabilities of new circuits containing greater numbers of devices, now VLSI, might not appear. Such fears seem to have vanished from the computer business, though not the perennial questions of which firms will get the largest share of the new markets.

Types of Computers

As pointed out in the earlier section on "System Design and the Microprocessor, " the essential elements of a small digital computer (exclusive of power supply and input/output devices) can be placed on one IC to create a single-chip microcomputer. A microprocessor is more limited in function, but when combined with the necessary memory and peripheral chips on a printed circuit board becomes a single-board microcomputer. From such products—selling for around $100 without cabinets and other auxiliaries—digital computers range upwards in size, speed, and cost to "supercomputers." Intended for complex scientific and technical calculations-e. g., modeling the Earth's atmosphere, designing airfoils or nuclear weapons—supercomputers are made by only a few manufacturers, and cost in the vicinity of $10 million each.

In between board-level microcomputers and supercomputers come a number of broad categories of machines: personal and small business computers like those made by Apple; minicomputers of various types; general-purpose mainframes. The latter, typified by many of IBM's larger models, can handle many different tasks at once. Table 3 outlines some of the conventional distinctions among these categories. The differences are not always clearcut and will blur even more as microcomputers become more powerful, computing power still cheaper, and computers of all types more versatile.

The central processing unit (CPU) for a simple computer—in fact, just a microprocessor—is shown schematically in figure 12. A microprocessor functions like the CPU of any computer—it brings information (in the form of binary bits) into an arithmetic logic unit, manipulating the bits in accordance with instructions

[18] L.M. Branscomb, "Electronics and Computers: An Overview, " *Science*, Feb. 12, 1982, p. 755.

Table 3.—Characteristics of Different Categories of Digital Computers

Family/Distinguishing features

Microcomputer: The central processing unit (CPU) consists of a microprocessor or single-chip microcomputer, sometimes several. The most common microcomputers use an 8-bit word and sell, without peripherals but otherwise complete and ready to operate, for under a thousand to a few thousand dollars. They will typically fit on a desk top—fig. 9—and do not require special training to operate. Examples include popular models sold by Apple and Radio Shack, along with the IBM Personal Computer,

Machines based on microprocessors or microcomputer chips with 16-bit word lengths are beginning to appear, particularly for business applications. These are nearer in performance to low-end minicomputers than to the &bit microcomputers originally developed for the hobbyist and personal computer markets.

Minicomputer: Microcomputers, by the definition above, could not have existed before the development of the microprocessor— i.e., before the early to mid-1970's. Minicomputers, in contrast, stem from the 1960's. A popular early mini introduced in 1965-the PDP-8, built by Digital Equipment Corp.—was the first low-cost, mass-produced computer of any type. It was designed around a 12-bit word and discrete transistors.

Minicomputers are small compared to mainframes, which can fill a room, As fig. 10 indicates, minicomputers are often about the size of a desk. Although not as portable as micros, many minicomputer models can be moved relatively easily within an office or factory environment.

Minis found much of their market as dedicated processors designed into more complex systems, or *in* specialized data processing tasks—e.g., industrial controllers, data acquisition systems for laboratory research, inventory management in factories. While such applications remain common, minicomputers are also widely used for general-purpose data processing. Often mainframes were needed in such applications only **a** few years ago; minicomputers tend to supplement rather than displace them.

At the lower end, it is increasingly difficult to distinguish minicomputers from the more powerful microcomputers. Many less expensive minis now rely on a single-chip processor. However, the smaller minicomputers typically use 16-bit words—e.g., the currently popular PDP-11 models made by Digital Equipment, or the Nova series of Data General. Larger, more powerful machines—sometimes called "superminis"--normally have a 32-bit word length. Examples of superminis are the Data General Eclipse series or the VAX models of Digital Equipment. In the 1960's, 32-bit words were found only in mainframes.

Most minicomputers carry prices in the $10,000 to $100,000 range. A principal distinction between minicomputers and mainframes is that minis seldom require either operators with a great deal of training or specially constructed facilities. Mainframes, in contrast, must usually be permanently installed; some large computers dissipate so much heat that air-conditioning is needed even in midwinter.

Mainframes: The CPU for a mainframe typically contains several thousand logic chips, usually bipolar for speed. Word lengths are commonly 32 to 64 bits. Mainframes often support multiple terminals and peripherals-fig. 11 —and generally require trained personnel onsite.

While IBM is the world's largest producer of mainframe computers, more than a dozen other firms build machines comparable in computing power. Mainframes can sell for $10 million or more–exclusive of peripherals–but the more popular general-purpose machines typically cost under $5 million.

SOURCE Office of Technology Assessment

from a program stored in memory, and sends information back to the memory or to an output device, The stored program, which made the modern digital computer possible by providing a means for telling the computer what to do without the need for hardware changes, accounts for a good deal of the information that enters and leaves the CPU. Even rather simple computer systems—figure 9—typically use several types of memory, which are described in more detail in appendix 3B.

Computers as Systems

In addition to CPU and memory, a computer system needs input and output devices (fig. ***13).*** In small computers, all the components may be integrated into a single self-contained unit. Peripherals such as disk and tape drives, terminals, and printers are made by large numbers of independent vendors, as well as by computer manufacturers. Nearly 90 American firms were producing terminals as of 1980, about half of these "smart" by virtue of embedded microprocessors, while nearly 30 had announced their intention to build 8-inch Winchester disk drives, a product just beginning to reach the marketplace at that time.[19] The market dynamics associated with the computer industry—rapid growth, intense competition, new entrants with new products—characterize peripherals and software as well as processors.

[19] "The Digital Age," *Electronics,* Apr. 17, 1980, p. 387; G. Slutsker, "28 Rivals Eye 8-Inch Disks But None Lands Big OEM Pact Yet, " Electronic News, Jan. 21, 1980, p. 40.

Figure 9.—Typical Microcomputer Intended for Personal and Small-Business Applications

Photo credit: Apple Computer, Inc

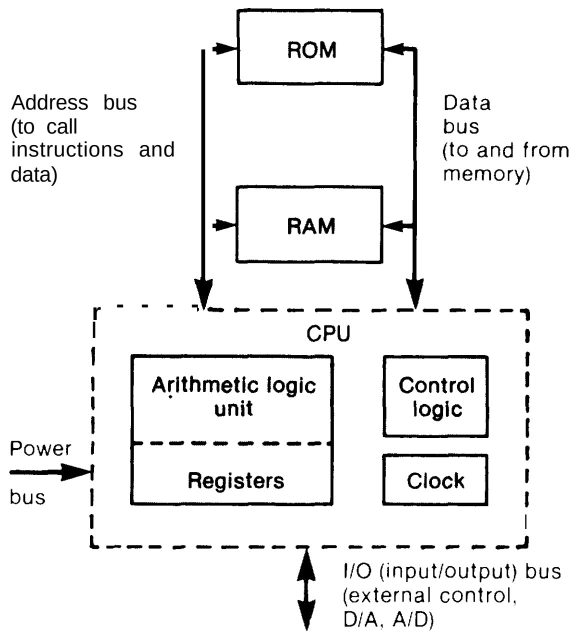
Figure 10.—Typical Minicomputer Installation Including a Pair of Terminals and a Printer

Photo credit: Digital Equipment Corp

Figure 11.— Data Processing Installation Built Around General-Purpose Mainframe Computer

Photo credit: Control Data Corp

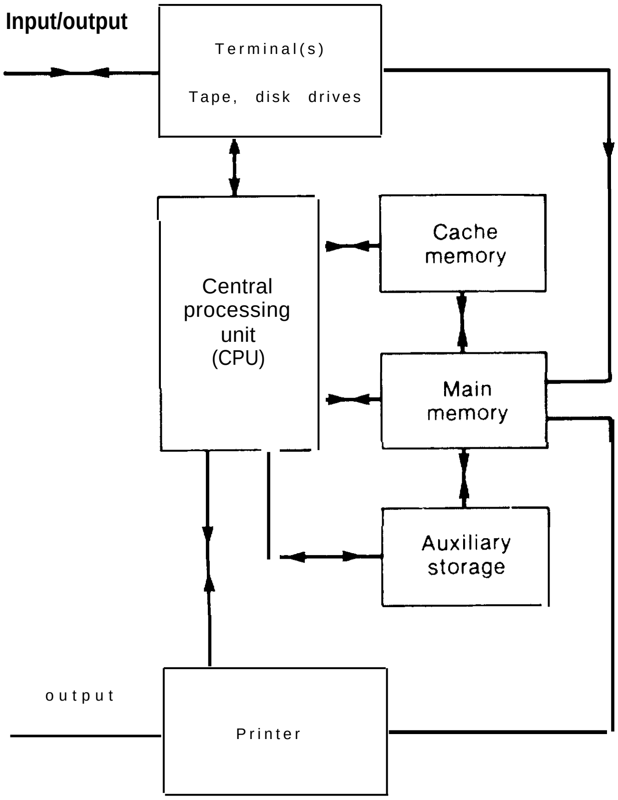
Figure 12.—Simplified Block Diagram of a Microprocessor System



D/A = Digital/analog conversion
A/D = Analog/digital conversion
ROM = Read-only memory
RAM = Random access memory

SOURCE Off Ice of Technology Assessment

Figure 13.— Elements of a General-Purpose Digital Computer System



SOURCE Off Ice of Technology Assessment

Technological change in computing has been rapid since the beginning of commercial production in the 1950's, but now the industry is perhaps facing the most comprehensive set of changes yet, These stem from "distributed intelligence," the dispersal of computing power to many farflung locations, In some respects, this trend began with the development of timesharing in the early 1960's. Time-sharing permits users at remote terminals to interact directly with a central processor, extending the capabilities of a powerful computer to many people simultaneously, It also uses the processor more efficiently, Even during big jobs the CPU may be idle much of the time; with timesharing, system software keeps the CPU busy by dividing its processing power among many people, each of whom is unaware of the others.

Conceptually, the next step beyond timesharing—for which each user needs only a "dumb" terminal (an input/output device with no function other than to communicate with the central processor)—is to link a central computer to satellite machines which can share the processing load. Many such distributed ***processing*** schemes are possible, among the more common being a mainframe supported by minicomputers. A mainframe or mini can also communicate with "smart" terminals that carry out limited computations, compile programs, and otherwise relieve the central or host computer of some of the work. Point-of-sale terminals found in retail stores often function as parts of distributed systems. The terminal not only acts as a cash register, but sends data on purchases to a central computer that can manage inventory, compare sales volume by brands, and provide other information to managers. Automatic banking machines are another familiar example; each automatic teller functions as a smart terminal linked to the bank's central computer(s). These systems may include hundreds of machines spread over several States.

Networking is a related term, referring to dispersed machines that communicate with one another but are each autonomous. Any one machine can transmit data to any other; control of the network maybe distributed over the system or may reside in a designated processor. In some but not all cases, networked computers not only communicate and share control, but also share the processing load. Local networks serve a limited group of users, such as a single office. At the other extreme, a multinational corporation might link computers located in many countries to form a worldwide network.

Computer Software

Physical equipment, or hardware—ranging from ICs, to disk drives, to networks—has been the primary subject above. But modern computers depend just as heavily on *software*. The programs that stand between user and CPU tell the hardware what to do. Arrayed in several levels, they range from applications software written in languages such as Fortran or Cobol —the only type of program that the typical user ever sees—to operating systems that supervise and coordinate both hardware and software elements. It is the software-architecture, operating system, compilers—that allows complex networks of computer and communications components to control steel mills, regulate air traffic, determine the path of a guided missile, distribute social security checks. Hardware and software in conjunction determine system performance, and customers weigh both aspects when making purchase decisions. In some cases this may entail buying software and hardware from different vendors and assembling a unique system. The spread of distributed intelligence, new applications of computers in homes and offices, shopfloor automation, computer-aided engineering analysis and design—all depend more heavily on versatile, reliable, user-friendly software than on hardware.

Since the beginnings of large-scale commercial production, computer hardware has become steadily cheaper relative to software. Costs for hardware have decreased by a factor of at least 1,000, holding processing power constant, over the past 25 years. [20] In marked contrast, software costs have not decreased appreciably, and may even have risen in real terms. A single line of programing, as a rule-of-thumb, costs in the range of $10 to $50—after inflation, about the same now as in 1955.[21] As a result, the largest part of the total cost to the user of a large computer system is now software, rather than hardware--figure 14. The chart applies to both purchased software-from computer manufacturers or independent vendors–and to user-developed programs; software maintenance is also included. At one time, many computer manufacturers provided system software such as control programs, language processors, and utilities free to hardware purchasers. Now, separate charges are the rule. For example, IBM currently sells about $1 billion worth of software per year, accounting for a little over 5 percent of the firm's total revenues; in newer systems such as the IBM 4300 series, nearly half the price of a typical installation is for software.[22] Similarly, more than half of the R&D commitment of a typical computer firm—measured either in terms of total expenditures or in terms of manpower—goes toward software.[23]

Cost trends for the development of software for dedicated applications—e.g., the logic for an embedded microprocessor—are similar. Even the simplest such application will require debugging and testing of the program to verify that it functions as desired. Software development for a microprocessor application may

[20] "Missing Computer Software," *Business Week*, Sept. 1, 1980, p. 46. The magnitude of the improvement depends on the type of system assumed.

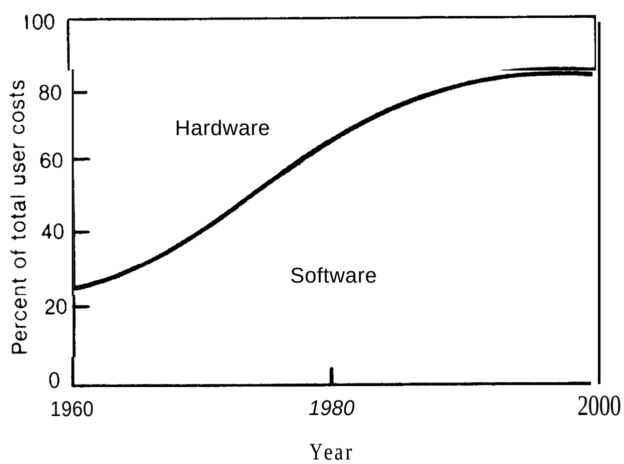
[21] While productivity in programing-as measured in lines of code per unit of time—has probably increased, the rate of increase has been orders of magnitude slower than for hardware performance. Birbaum, for instance, points out that while programer productivity has increased by about a factor of 3 since 1955, system performance-to-cost ratios have gone up by roughly 10^5 over the same time period—J. S. Birbaum, "Computers: A Survey of Trends and Limitations," *Science*, Feb. 12, 1982, p. 760.

In some programing tasks, productivity has probably remained rather stable—perhaps even decreased. Applications programing may now be somewhat more efficient because of improvements in higher level languages. Productivity in systems programing, or developing software for dedicated microprocessors, microcomputers, and minicomputers, has probably not improved as rapidly; when systems become more complicated, many of the stages in program development—from conceptual design to debugging—become more arduous. Even a relatively simple program may have of the order of 10^{20} different execution paths, depending on the number of loops, branches, and subroutines. Costs per line can escalate as program size and complexity increase. Another common rule-of-thumb is that a man-month of effort is required to demonstrate that 100 lines of code is, for practical purposes, error-free and functionally correct.

[22] "Missing Computer Software," op. cit.

[23] "Computer Technology Shifts Emphasis to Software: A Special Report," *Electronics*, May 8, 1980, p. 142.

Figure 14.—Relative Hardware and Software Costs Faced by Users of Larger Computer Systems



SOURCE: Office of Technology Assessment.

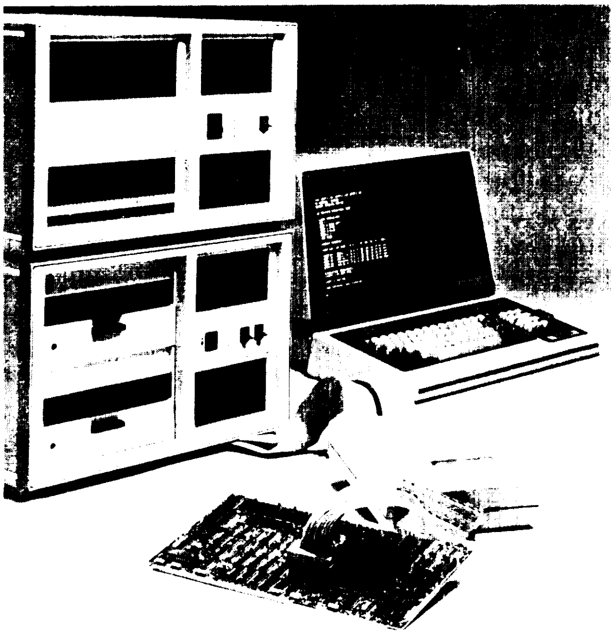


Photo credit Tektronix, Inc

Microprocessor development system

cost several hundred million dollars, with estimates for ***1985*** running to ***$3*** million or more .24

The rising relative costs of software have been one factor in the rapid growth of independent firms that develop and market programs of all types. Many computer manufacturers have traditionally been rather hardware-oriented, leaving an attractive market for vendors who concentrate on software (see app. C, "Computers: A Machine for Smaller Businesses," on the role of "systems houses" in the development of the minicomputer market). Even IBM—which has built its market dominance in larger machines on software as much as hardware—has turned to independent software firms to supply programs for its personal computer. Independent software vendors sell perhaps $1 billion in off-the-shelf programs per year, and twice that amount in custom programing.[25]

Growth of Computing Power

Figure 14 gave one picture of the rapidity of change in computer technology, and in the computer industry in general. But change extends far beyond the relative costs of hardware and software, the rapid growth in the microcomputer market (now about 50 percent per year), or continued increase in performance cost ratios for computer systems. And, while distributed intelligence may eventually have broader and deeper effects on the way people live and work than big machines, the absolute rise in computing power delineated in table 4 illustrates simply but dramatically how rapidly the capabilities of the most powerful digital computers have increased—nine orders of magnitude since the close of the Second World War, six orders of magnitude in the 30 years since the introduction of the first commercial machine, the Univac I. All the computers listed in table 4 would be classed as mainframes, and those of recent years as supercomputers—representing the maximum in computing power available at a given time. *

While the biggest computers have been growing in speed, smaller machines--like all computers—have been growing in performance per dollar. Table 5 compares an 8-bit single-board microcomputer representative of 1970's technology to the IBM 650—a first-generation vacuum tube processor of the mid-1950's. The two machines are roughly comparable in computing power, but the modern microcomputer is orders of magnitude smaller and cheaper,

[24]"Missing Computer Software, " op. cit. One supplier of microprocessors and microcomputers has estimated that a typical mid-l 970's application carried a software development cost of about $20, 000 ($20 per 1 ine of code), but by 1980 the cost was $100,000 to nearly half a million dollars ($35 per line of code, but also many more lines], Meanwhile the hardware costs have remained about the same—in the vic init y of $100 per unit, See J G.Posa, "Intel Takes Aim at the '80s, " *Electronics*, Feb. 28, 1980, p. 89.

'SW. D. Gardner, "The Key to Greater Productivity," *Dun's* Review, August 1980, p. 74. The total value of computer software in use worldwide probably exceeds $200 billion.

*Arithmetic operations per second, the measure used in the table for comparing computing power, is not a perfect yardstick because many data-processing programs are limited by operations other than arithmetic— .g., inverting matrices, More sophisticated comparisons employ "benchmark" programs based on representative tasks. Arithmetic operations as used in the table have the advantage of being easy to understand and applicable to early model computers, some of which could not execute modern benchmarking programs.

Table 4.—increase in Computing Power Over Time

Year	Model	Computational speed (arithmetic operations per second)
1944	Harvard Mark I (electromechanical)	0.4
1946	Eniac	45
1951	Univac I	270
1953	IBM 701	615
1961	IBM 7074	33,700
1963	CDC 3600	156,000
1965	IBM 360/75	1,440,000
1972	CDC Cyber 176	9,100,000
1976	Cray 1	80,000,000
1981	CDC Cyber 205	800,000,000

SOURCES J R Bright, "Technology Forecasting Literature Emergence and Impact on Technological Innovation, " P Kelly and M Kranzberg (eds), *Technological Innovation: A Critical Review of Current Knowledge* (San Francisco. San Francisco Press, 1978) p 300, "The Digital Age," *Electronics*, Apr 17, 1980 p 382, P J Schuyten, "The Battle in Supercomputers " New *York Times*. July 22, 1980, p D1

and—at least as significant—vastly more reliable.[26] Note that these computers are separated in time by only two decades. Figure 15 gives an alternative picture of growth in performance per dollar. The plot shows the decline in price for a minicomputer family—after 1965 the pioneering PDP-8, although the first several years apply to an earlier model–the rapid fall stemming in part from learning curve phenomena as for semiconductor devices. Drops in prices for the semiconductors a machine contains—figure 16—also lead to cost reductions. Digital Equipment Corp. introduced the PDP-8 at $18,000; by the early 1970's some versions were priced as low as $2,500.[27]

"Generations" of computers can be distinguished based on advances in the technology. For example, the IBM 650 (table 5) represents a first-generation machine, the F-8 microcomputer third generation. Zeroth-generation systems were similar to the 650 in using vacuum tubes, but early computers such as Eniac lacked the ability to execute stored programs, the hallmark of the modern digital computer. To change a program in Eniac meant altering

[26] Integrated circuits typically exhib it reliabilities—measured as mean times between failure (ch. 6)—of the order of 10^{11} hours/gate. Thus, a typical microprocessor containing 10,000 gates might have a mean time between failures of about 1 () million hours, or 1,000 years. In contrast, mean times between failures for discrete transistors are about 10^{8} hours, for vacuum tubes, less than 10^{6} hours, See, S.Middelhoek, J.B. Angell, and D.J. W.Noorlag, 'microprocessors Get Integrated Sensors, " IEEE *Spectrum*, February 1980, p. 42.

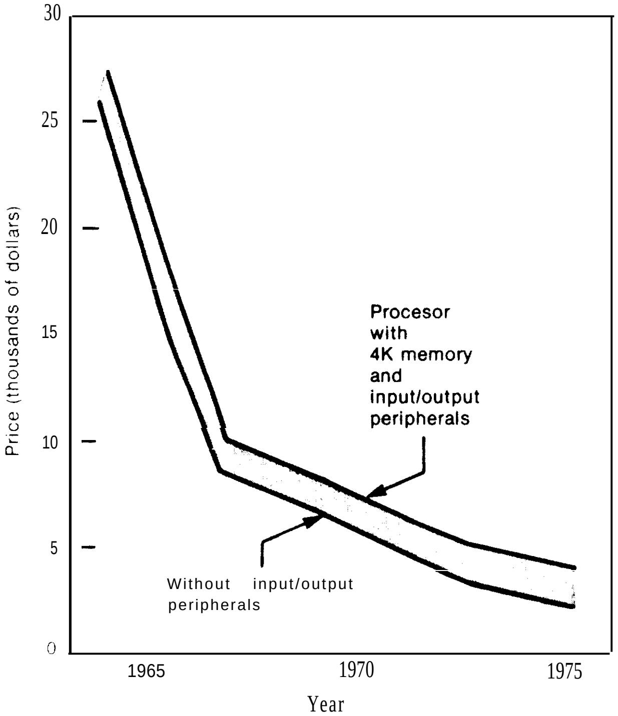
[27] G. Lewis, "Small Computers, " *Electronic News*, Jan. 25, 1982, sec. I I, p. 70.

Table 5.—Comparison of IBM 650 (1955) and Fairchild F-8 Microcomputer (1970's)

	IBM 650	F-8	Remarks
Physical volume (ft^3)	270	0.01	F-8 about 30,000 times smaller
Weight (pounds)	5,650	1	
Power consumption (waits)	17,700	2.5	F-8 consumes 7,000 times less power
Memory (bits)	3K main, 100K secondary	16K ROM, 8K RAM	
CPU	2,000 vacuum tubes	20,000 transistors	650 also needed many discrete resistors and capacitors
Time for adding two numbers (microseconds).	750	150	
Reliability (mean time between failures)	Hours	Years (3 million to 10 million hours is a typical mean time between failures for a current microprocessor—more than 300 years—but the subsystems with which the microprocessor communicates—e.g., terminals, printers—may be much less reliable)	F-8 at least 10,000 times more reliable
cost	$200,000 (1955 dollars)	Under $1,000 with terminal	

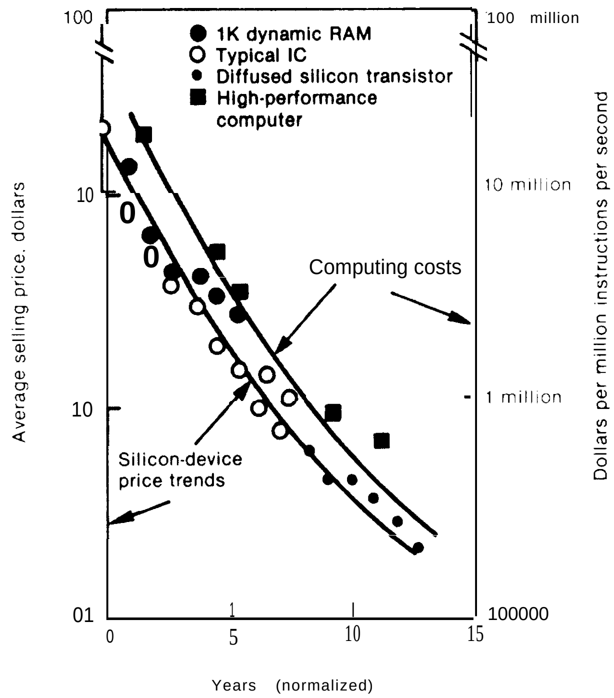
SOURCES IBM 650 information from "1978 First Quarter and Shareholders Meeting Report, Texas Instruments, Inc Fairchild F-8 information from" J G Linvill and C L Hogan, "intellectual and Economic Fuel for the Electronics Revolution, " *Science*. Mar 18, 1977, p 1107

Figure 15.— Minicomputer Price Trends (12-bit Digital Equipment Corp. models, PDP-8s after 1985)



SOURCE C G Bell J C Mudge, and J E McNamara, *Computer Engineering* (Bedford Mass Digital Press, 1978), p 194

Figure 16.— Parallel Decreases Illustrate How Costs of Computers Depend on Costs of Semiconductors



SOURCE I M Mackintosh, "Large-Scale Integration Intercontinental Aspects," *IEEE Spectrum,* June 1978 p 53

thousands of patchcords—in essence changing the hardware—a task that could take several days.

Current machines represent generation three-and-a-half or four, table 6, although the notion of generations has lost much of its meaning with the rise of distributed processing, A good deal of attention has recently focused on planning within Japan for fifth-generation technology—which, as the table indicates, might be characterized by major advances in software associated in a general way with "artificial intelligence." Among the goals of Japan's fifth-generation computer project are input and output in natural language—as ordinarily spoken or written. As the table indicates, the next generation may also be identified with new methods for communicating among computers, and perhaps some replacement of silicon ICs with higher speed devices. Hardware developments such as Josephson junctions would make natural language programing and related developments such as voice recognition easier to achieve by speeding the processing of the very complex algorithms required.

As noted in the table, the current generation is one of specialization, characterized by continuous rapid development on many fronts: minicomputers, microprocessors and microcomputers, distributed processing, new programing languages, improved peripherals, special-purpose machines such as array processors. As a result, it makes less sense to speak of "generations. " Computing power is becoming so widespread and pervasive that to focus on the characteristics that different systems have in common may obscure the true significance of specialization, and the distribution of machine intelligence to new applications—many of them quite different from those originally associated with "computers. "

Table 6.-Characteristics of Generations of Computer Technology

Generation	Period	Representative models	Description	Typical applications
Zero	1940's	Eniac	No stored program capability; vacuum tube processor.	Preparation of ballistics tables.
One	Early 1950's	Univac I	Stored program, but in binary machine language only. Vacuum tube processor.	Scientific and technical calculations (aerodynamics, nuclear weapons design) business (accounting, Inventories, payrolls); Government (census).
Two	Late 1950's	IBM 7090	Higher level languages such as Cobol; CPU uses discrete transistors. Magnetic core memory common, along with line printers for output. Punched cards used for data entry.	As above, but much more widespread.
Three	1960's	IBM 360 series; Burroughs 6500	Hybrid ICs—combining discrete transistors and integrated circuits on a single substrate (IBM 360 series) or small-scale ICs (Burroughs 6500 and others) used in CPU. Time-sharing available.	Continuing spread of data-processing applications as costs decrease. Real-time processing becomes more common.
Three-and-a-half or four	Late 1960's to present	IBM 370 series; DEC PDP-11	Large-scale ICs; distributed processing, networking. Proliferation of special purpose computers; rapid growth of minicomputer markets. Microcomputers developed.	Great Increase in specialized, dedicated applications, particularly for minicomputers. Data base management systems spread. Networking and distributed processing point toward merging of data processing and data communications, typical applications being electronic funds transfer, Microprocessors make many products "smart, " as well as substituting for custom logic,
Five	Late 1960's or 1990's	?	Natural language programing; voice recognition, speech synthesis. Gallium arsenide ICs or Josephson junction devices may replace silicon ICs in CPUs.	

SOURCE: Office of Technology Assessment. See, in general, S Rosen, "Electronic Computers: A Historical Survey, " *Computing Surveys*, vol 1, 1969, p 1, J T Soma, *The Computer Industry" An Economic-Legal Analysis of Its Technology and Growth* (Lexington, Mass. Lexington Books, 1976), pp 9-30, "Digital Computers: History," *Encyclopedia of Computer Science*, A Ralston and C.L. Meek (eds) (New York: Petrocelli/Charter, 1976), pp 474-495.

Applications of Computers

As pointed out earlier, the microprocessor was originally developed as an alternative to custom ICs for a line of hand calculators—an example of a dedicated or embedded application where the computer is invisible to the user, Such applications of small processors far outnumber generalized data processing. In most dedicated systems, the programs are permanently stored, and the user interacts with the machine through switches, control knobs, or—as in the case of a word processor—a keyboard. A large commercial aircraft may contain a dozen or more computers, but the pilot need never know of their presence. His interfaces are instruments and flight controls. In the same way, when a minicomputer or mainframe supports word processing applications, the typist may never see or be aware of the computer. As this example shows, dedicated applications need not be restricted to small machines. Furthermore, the distinction between dedicated and general-purpose applications is not always clear-cut. A mainframe computer might support dozens of word processing stations, while at the same time running data processing programs in both batch and time-sharing modes.

Table 7 illustrates something of the range of current applications of computing power, while an example from the field of industrial process control is described in more detail in appendix 3C. Note that even in the rather arbitrarily defined data processing category, several of the familiar examples—airline reservations and tickets, point-of-sale terminals—depend on dedicated machines. Also note that dedicated applications help make the data-processing systems themselves function; controllers for disk or tape drives are often based on microprocessors.

Leaving aside the overlaps among categories—because so many technologies blur and merge as intelligence is added—the breadth of

Table 7.—Typical Applications of Computing Capability

Example	Usual Type of Computer
Data processing	
Business records (accounting, payroll, order processing and billing, production control, inventories, taxes, banking).	Mainframe, mini, or micro, depending on size of business.
Government records and statistics (census and other data bases, tax records, social security, economic data).	Mainframe.
Scientific and technical (social science data bases, engineering calculations, modeling of complex systems).	Mainframes for batch and interactive processing; micros and minis for laboratory automation as well as specialized applications such as modeling chemical reactions.
Medical records.	Mini or mainframe.
Airline reservations.	Mini or mainframe.
Point-of-sale terminals, electronic cash registers.	Micro, but may be part of distributed system.
Communications and control	
Multiplexing and transmission of voice and alphanumeric data.	Varies.
Telephone exchanges.	Mainframes.
Private exchanges (PBX, PABX).	Micros and minis.
Facsimile transmission.	Minis and micros.
Teletext, viewdata.	Micros.
Air traffic control.	Mainframes.
Military systems	
Signal processing (radar, sonar).	Mainframe or mini, depending on need for portability.
Navigation	As above, or micros.
Fire control.	As above, or micros.
Flight control.	Micros.
Industrial systems	
Batch process control (machine tools, assembly robots, heat treating, materials handling, steelmaking, typesetting).	Minis and micros.
Continuous process control (petroleum refining, rubber and synthetic fibers, basic chemicals, paper products, foods).	Mainframes and minis.
Computer-aided design.	Mainframes and minis.
Energy production, conservation and control (turbine startup, electric utility load management, process heat, building heating, ventilation, and air-conditioning).	Varies.
Environmental monitoring and pollution control.	Minis and micros.
Education and training (computer-assisted instruction).	Varies.
Measurement and testing (medical diagnostics, nondestructive inspection, chemical analysis).	Minis and micros.
Office automation	
Word processors.	Micros and minis
Copiers.	Micros.
Calculators and accounting machines.	Micros.
Consumer products	
Automobiles (engine control, driver information, diagnostics).	Micros.
Home entertainment (electronic and video games, personal computers).	Micros.
Appliances (refrigerators, microwave ovens, sewing machines).	Micros.
Thermostats and environmental controls.	Micros.
Calculators.	Micros.
Cameras.	Micros.
Electronic watches.	Micros.

SOURCE Office of Technology Assessment

applications is striking. In fact, it is difficult to think of manufactured products or processes that could not use computing power in some form. In a recent 35mm camera design, the number of mechanical parts dropped from nearly 1,300 to 900 when a single-chip microcomputer was incorporated.[28] Some observers have predicted that a typical home will contain a dozen or more computers by the 1990's

[28] "Canon's Fujio Matarai: Strategies for the U. S. Market," *World Business Weekly*, Oct. 19, 1981, p. 20.

—a proliferation often compared to that of the fractional horsepower electric motor, Limitations on such new and specialized applications often stem from software engineering problems or total system hardware cost—but seldom the cost of the computing power itself. For example, to extend the use of microprocessors in automobiles to nonskid braking systems is possible, even straightforward, from an engineering standpoint, but nonetheless expensive—primarily because of the cost of the sensors and actuators required. On the other hand, the design of a practical collision avoidance system for ordinary driving is still limited by engineering problems that, broadly speaking, can be considered software. It is quite difficult to develop algorithms for unambiguously detecting collision hazards, and for adapting the output of a collision hazard identification system to the controls facing the human operator—steering wheel, accelerator, brakes.

Summary and Conclusions

Electronics technology—used for transmitting and manipulating *information* via electrical signals-has been evolving from analog toward digital, driven in the broadest sense by applications of computing power. Digital communications have advantages over analog, and as both computing and communications have moved toward digital technologies, the semiconductor industry has been called on to provide new kinds of building blocks, Even in the consumer electronics industry, digital equipment is being developed.

Although new products such as video cassette recorders and video disks have reached the marketplace in recent years, consumer electronics technologies move slowly compared to semiconductors or computers. Commercial viability depends on consumer appeal, which can come from technology but also from many other sources. The success of the Sony Walkman—a personal audio tape player—is more the result of a good match between the marketplace and the engineering development laboratory than of new technology. Coming generations of consumer electronics products—e.g., integrated home entertainment systems—-will continue to stand or fall on product design and marketing, with technology as only one dimension.

Still, the products of the consumer electronics industry have been transformed by semiconductor devices. Beginning with the transistor radio—and extending to the digital technology embodied in video games, pocket calculators, and electronic watches—products that people see and use every day are practical only because of semiconductors. As microelectronics technology has progressed from discrete transistors to small-scale ICs and then very large-scale circuits, new applications in electronic systems of all types have emerged. The growth of the computer industry has followed advances in semiconductors, as have military applications ranging from missile guidance to war gaming,

VLSI has brought forth the "system-on-a-chip"--a prime example being the single-chip microcomputer. While one IC can now accomplish more than a room full of computing equipment three decades ago, microprocessors and microcomputers are also used in large numbers for quite different purposes—namely for replacing "hard-wired" logic, Engineers can trade off the hardware costs of custom design and manufacture against the software costs of developing programs to be permanently stored in the memory associated with a dedicated microprocessor or microcomputer. The programed logic can then control a microwave oven or fly an airplane.

With advances in density and performance have come higher costs for IC design, and for the sophisticated processing equipment needed to make VLSI devices. These have driven the capital requirements for entering the semiconductor industry upward; firms striving to compete at the leading edge of the technology now depend on both computer-aided circuit design and computer-aided process control. In semiconductor processing, as in the development of computer software, fundamental understanding from the viewpoint of engineering science has lagged practice. Semiconductor fabrication is an art, as is programing, Putting these two technologies on firmer underpinnings will be critical for avoiding future bottlenecks in the development of the semiconductor and computer industries.

As digital computers evolved they grew bigger in processing power, smaller in size. The range in size and capability of digital computers is now truly awesome—from single-chip microcomputers costing *$100,* to supercomputers like the Cray-1 that operate at speeds limited in a very real sense by that of light (because electrical signals propagate at speeds that can be no greater, and the time to move signals within the processor limits computational speed). At the same time, the seemingly mundane problem of transferring the heat dissipated in the chips out of the system is one of the critical elements in the design of a high-performance machine.

Applications of computers have sometimes been driven by the availability of the technology, and its continually decreasing costs, sometimes by newly recognized needs—forces that interact continuously. A great deal of technical ferment is presently centered on software of all types, Markets for small computers—intended for personal use as well as applications in business and industry—are expanding rapidly. The most pervasive trend is the widespread distribution of intelligence to the points where needed. Machines of all types are getting smarter and more specialized, and computers (including microprocessors) already outnumber people in industrialized countries like the United States.

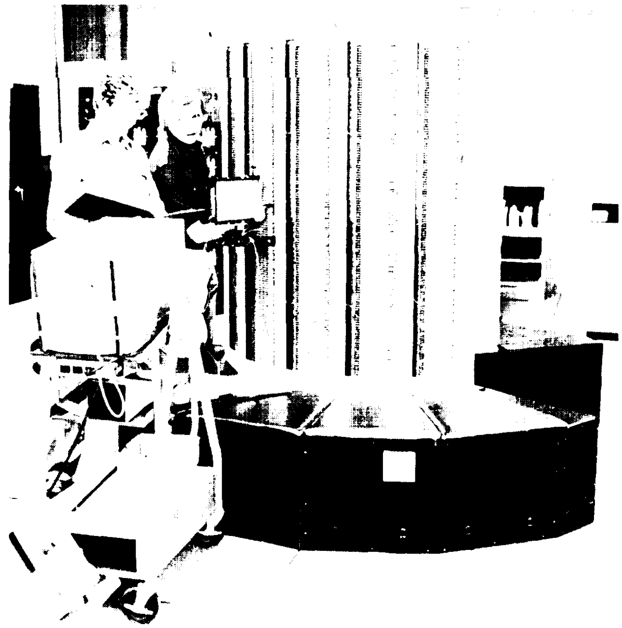


Photo credit Cray Research

Central processing unit for a supercomputer

Appendix 3A.—Integrated Circuit Technology

Types of Integrated Circuits

Table 3A-1 outlines some of the principal varieties of ICs.

As table 3A-1 indicates, ICs come in two major varieties: MOS and bipolar. Many small- and medium-scale circuits are bipolar, as are some LSI chips, but over the course of the 1970's MOS technology became dominant for LSI. Bipolar chips are less dense than MOS, typically by factors of about four—i.e., transistors and other circuit elements cannot be located as close together; they also consume more power—one of the reasons that the circuit elements cannot be packed more closely is the need for heat dissipation. The greater number of circuit elements that can be placed on a given area of silicon using MOS technology often ***leads to*** significantly lower costs. By contrast, the chief advantage of bipolar technology has been high operating speed. However, the speeds of MOS ICs have been improving more rapidly than the speeds of bipolar devices. Where high speed is critical—as in

Table 3A-1.—Common Terminology and Classifications of Integrated Circuits

Assembly

Monolithic--all circuit elements fabricated on a single semiconducting substrate (usually silicon).

Film—conducting layers are deposited on an insulating substrate to form the circuit.

Hybrid—combines several ICs and/or discrete transistors in a single package, often using film technology for components and interconnections.

Input-output Characteristics

Analog (also called linear)—levels (e.g., voltages) of input and output signals vary continuously over a range.

Digital—input and output signals have values limited to either of a pair of nominally discrete values (e.g., voltage levels of O or +5 volts).

Transistor technology

Bipolar—conduction takes place through motion of both electrons and holes (a hole is an electron vacancy—the absence of an electron where one would ordinarily be; the same current can be carried by electrons moving in one direction, by holes moving in the *opposite* direction, or by a combination of electron and hole motion); *control* of the transistor is through a *current* signal.

MOS (metal -oxide-semi conductor)— MOS transistors are *unipolar* rather than bipolar; conduction is by either electrons or holes, but not both, and the transistor is controlled by an impressed voltage (MOS transistors are actually subsets of the broader class of field effect transistors (FETs), but the MOSFET is by far the most common type of FET).

Digital device applications

Logic** circuits*— both bipolar and MOS ICs are used for digital logic circuits; the logical operations are performed by arrays of ***gates, each of which implements a Boolean function such as AND, OR, NOR, NAND; the gates themselves consist of groups of circuit elements typically including one or more transistors plus resistors and capacitors.

Bipolar transistors can be grouped by *logic families* such as TTL (transistor-transistor-logic) and I^2L (integrated-injection-logic), the names of which characterize the gate circuitry.

MOS circuits are classed somewhat differently; n-MOS, the most common, refers to circuits using “n-channel” MOSFETs, where the current is carried by electrons, p-MOS to “p-channel” technologies where the current is carried by holes, and c-MOS to “complementary” MOS where n- and p-channels coexist in the same IC.

Microprocessors—digital logic circuits that can serve as processing units for digital computers—i.e., can execute programs; most microprocessors are MOS ICs.

Microcomputers— ICs that contain a processing unit plus memory circuitry.

Memory--ICs that can store digital data in an array of logic gates, each storage location containing a “bit” of binary (“O” or”1”) information; the number of bits stored in a single IC presently ranges up to more than 64,000 (in a 64K random access memory (RAM)). (See table 3A-2 for more detail on memory circuits. Some types already have higher capacities than the 64K RAM.) Most memory chips are MOS ICs.

Circuit density

Levels of integration, or packing density, for ICs are grouped by order of magnitude of the number of devices on a chip.

Discrete—single individual y packaged active device—e.g., a transistor.

***Small-scale** integration* (SSI)—refers to ICs with of the order of 10 active devices.

Medium-scale integration (MSI)— ICs containing of the order of 100 devices. The simplest gates used in dynamic random access memory (RAM) chips consist of one transistor plus one capacitor, and IC density is sometimes referred to in terms of gates per chip rather than devices per chip. Because some gate designs use several active elements, devices per chip is more meaningful.

Large-scale integration (LSI)–ICs with roughly 1,000 to 10,000 or more devices. A so-called 1 K RAM can store 2^{10} or 1,024 bits of digital information. Each memory cell includes at least one gate. Thus, a 1 K RAM, which needs other devices for getting the bits into and out of the chip, is an LSI circuit. So is a 4K RAM, which includes about 5,000 devices. 16K RAMS (close to 20,000 devices) are usually considered LSI circuits, with 64K RAMs (2^{16} or 65,536 memory cells, plus several thousand devices for getting the bits into and out of the chip) the lower end of very large-scale integration.

Very *large-scale integration (VLSI)—ICs* containing of the order of 100,000 devices. In addition to 64K RAMS-which entered mass production during the late 1970’s—other VLSI devices include 16-bit microprocessors such as the Motorola 68000, which contains about 69,000 devices. By 1982, the densest circuit produced was a microprocessor built by Hewlett-Packard with 450,000 devices. Pilot production of next generation RAMs—256K chips—began during 1983. ICs containing on the order of a million devices will probably be given a name such as ultra large-scale integration.

SOURCE: Office of Technology Assessment

processors for large computers—bipolar ICs remain the technology of choice. For consumer products such as calculators and watches, for primary read/write memory in computer systems, and in other applications where cost is more important than speed, MOS is generally specified. Some of the distinctions between bipolar and MOS technologies may blur and disappear as IC technology continues to advance, with much of the impetus for such developments likely to come from R&D efforts focused on improvements in MOS.

A principal application of bipolar chips is digital logic; bipolar logic circuits are often used in conjunction with MOS microprocessors and memory circuits as parts of complex systems including many ICs. There is considerable demand for bipolar small- and medium-scale devices that can serve as universal building blocks; here, their low packing density is not necessarily a handicap. The great majority (perhaps 90 percent) of standard bipolar logic circuits belong to the TTL—or transistor-transistor logic—family (table 3A-l). Typical examples of small- and medium-scale TTL circuits would be counters, buffers, and digital/analog (D/A) or analog/digital (A/D] converters. Buffers store strings of binary data for short periods of time; they are used to adjust and coordinate data rates between different parts of a system. D/A and A/D converters serve as interfaces with analog components that provide input signals to the system or receive its output. A common output device is a cathode-ray tube, which must be driven by an analog signal and thus depends on D/A converters when fed digital information from a computer or word processor. A/D and D/A converters are also used in conjunction with many types of sensors and actuators (a familiar example of a sensor is a thermometer; an electric motor can function as an actuator).

How Microprocessors Work

A typical microprocessor is an MOS IC including an arithmetic logic unit, several registers, control logic, and paths for moving data among these (fig. 12). Within the arithmetic logic unit, groups of binary bits called words (or bytes) are added and subtracted by moving and manipulating the strings of bits among the registers. These operations are all performed on numbers represented in binary form; the conventional decimal (base 10) number 6, for example, is written in binary (base 2) as 0110.

In a 4-bit microprocessor, the standard word consists of 4 binary bits; an 8-bit microprocessor has a word length of 8 bits. In general, the longer the word, the faster and more powerful the microprocessor—but also the higher the costs for programing and system development, as well as for production or purchase. Four-bit microprocessors are suited to inexpensive pocket calculators or controlling a simple system like a microwave oven. Microprocessors with longer word lengths are used for more demanding applications. The most complex microprocessors now in common use have 16-bit words; 32-bit chips will follow. Large mainframe computers such as those produced by IBM are typically designed around words having 32 or more bits.

The control logic in a microprocessor [fig. 12) regulates the flow of binary information within the chip. Many microprocessors include a clock that synchronizes the operations performed, although sometimes a separate clock chip must be provided. The faster the clock speed, the faster the microprocessor can manipulate information, everything else equal, Typical 8-bit microprocessors operate at clock speeds in the range of 5 MHZ—5 million cycles per second—which does not mean that they can perform computations at this rate. The various instructions that the microprocessor carries out normally take several clock cycles, and the logical operations that these implement depend on the instruction set—more generally, on the architecture of the processor. These features of the design determine the permissible ways that binary data can be manipulated. The tradeoffs involved in defining the architecture and instruction set for a microprocessor mean that some microprocessors perform certain kinds of tasks faster or with simpler programing than others. But because all the binary operations performed in the arithmetic logic unit are primitive, a microprocessor with a clock rate of 5 or 10 MHz executes functions such as subtraction or multiplication at rates which are only a small fraction of this.

The microprocessor must also be able to pass binary information back and forth to chips that provide memory, A/D or D/A conversion, and a variety of specialized functions. This is done through input/output (1/0) ports, The circuit paths along which the bits travel are called buses (fig, 12). As implied, a microprocessor cannot function by itself, but must be supported by other circuitry. As a minimum, the microprocessor has to communicate with a ***memory*** sufficiently large to hold the program being executed, (A ***single-chip microcomputer*** includes on-chip memory for program storage, in contrast to a microprocessor, which does not have built-in memory.) The program itself consists of a set of instructions, coded in binary form, which tell the processor how to manipulate the bits in its registers, The processor uses an address to fetch the appropriate information from memory or to send information back to memory locations. Likewise, 1/0 buses and ports have associated addresses. Many of the operations performed by the processor are simply matters of getting the string of bits into or out of the registers.

In addition to memory, a microprocessor must be connected to a power supply, and it often communicates with external devices—e.g., ***sensors or transducers*** that generate electrical signals corresponding to the magnitude of parameters such as temperature, position, or pressure. Usually the transducer output is an analog signal—for example, a resistance thermometer produces a continuously variable voltage—and an A/D converter must be interposed. This can be built into the transducer–some microcomputers are also available with built-in A / D and D / A converters—but is usually an independent circuit. A / D converters (and D / A) tend to be limited in speed and precision; they are also expensive—a single converter may cost more than the microprocessor it is used with.

Peripherals are information-handling equipment external to the central computer, whether it be a microprocessor or a larger machine. Examples of peripherals are bulk memory (in the form of arrays of IC chips, or magnetic disks and tapes), terminals (typically keyboards with or without CRT display screens), and printers. *Interfaces* are generally needed to allow a microprocessor or computer to communicate with the peripherals. While in the past, most interface circuits were custom-designed from SSI and MSI chips, LSI parts are now available in standard form to implement common functions such as interfacing with and controlling a CRT terminal. Typically, interfacing is much more demanding from a software than from a hardware standpoint.

Memory Circuits

Microelectronic devices–generally ICs–that store information in binary form come in many varieties, as outlined in table 3A-2. Declining prices for RAMs, in particular, have led to rapid increases in demand for applications in computer-based systems of all types. As specialized memory chips—for instance, erasable-PROMs—become cheaper and easier to use, still more applications for microprocessor systems will open.

How Integrated Circuits Are Made

Circuit fabrication begins with a silicon wafer sawn from a carefully grown cylindrical crystal. The wafer itself is a fragile disk, as thin as 0,010 inches, on which hundreds of ICs will be simultaneously created before being cut apart into individual chips. Although a few large manufacturers grow their own crystals, most semiconductor-quality silicon is produced by independent firms. The composition—particularly the oxygen content—must be carefully controlled, as must the flatness of the wafers, which is critical for high yields in the subsequent lithographic processing. At present, the most common wafer diameter is 4 inches. Bigger wafers reduce production costs because more circuits can be made at once; therefore, as IC fabrication technology advances, wafer diameters tend to grow.

Table 3A-2.—Principal Types of IC Memory Circuits

Designation/Function
Read-only memory (ROM): Contents are permanently stored during manufacturing; memory can thereafter be read but not altered. Commonly used for program storage in microprocessor-based systems, the memory contents in a ROM are normally determined by the masking patterns used in fabricating the circuit (IC manufacture is described in the next section).
Read-write memory (RWM): Memory contents can be written over and changed, as well as read. Applications include storage of data, output, programs, and other general memory requirements.
Random-access memory ***(RAM):*** Common name for IC read-write memory chips. Strictly speaking, random-access means only that any particular memory cell can be addressed directly and the contents retrieved. By this meaning, ROM chips, for example, are also random-access. Nonetheless, in common usage the term random-access or RAM now applies only to read-write memory. In contrast to IC RAMs, a bubble-memory device is not random access because data is stored in a string of magnetic bubbles which can only be read or written sequentially by passing the string through a detector until the desired address is located. The time to access any memory location in a RAM is nominally the same; in a serial device such as a bubble memory, the time depends on where in the string the memory location happens to fall with respect to the detector. RAM circuits can be static or *dynamic.* The basic difference is that dynamic RAMs store data in memory cells that rely on capacitance. As the charge gradually leaks off capacitors, they must be "refreshed" several times a second by a voltage pulse. (Many microprocessors provide built-in refresh capability because the need is so common.) Static RAMs, in contrast: do not require refreshing; each cell will retain its contents as long as power is supplied to the chip. Both static and dynamic RAMs are ***volatile*** memory devices. This means that their contents are lost when electrical power is removed. In contrast, magnetic tapes or disks are *nonvolatile* because they retain the data stored whether or not supplied with power. Bubble-memory devices are nonvolatile, as are ROMS and magnetic core memory.
Programmable ***read-only memory (PROM):*** ROMs with memory cells in which data can be stored after manufacture are called—in contrast to a ROM—PROMS. Some PROMS are permanently programed by a process analogous to blowing fuses, after which the contents of the memory cells cannot be altered. In other types, the contents of the cells can be erased—e.g., by exposing them to ultraviolet light—and then rewritten. PROMS are widely used in *system development—i.e.,* preparing the software on which a microprocessor-based system functions—as well as in low and medium volume production applications. In a typical development project, software is stored in PROMS for testing and debugging. Once the software functions properly, and the system is ready for production, programs might continue to be stored in PROMS, which would be programed during the manufacture of the system. Alternatively, they could be transferred to ROMs. The choice between PROM and ROM is a matter of manufacturing cost. PROMS are generally cheaper in small volume production, ROMs at high volumes.

SOURCE Off Ice of Technology Assessment

Limitations centered around wafer flatness and lithographic capability prevent more rapid movement toward large wafer diameters.

After polishing, an epitaxial layer is sometimes grown on the wafers, followed by heating in a furnace to produce a layer of silicon dioxide, Oxidation produces the insulating layers that separate various parts of the circuit from one another, Portions of the oxide layer are later selectively etched away, leaving windows open to the silicon beneath.[1] As well as isolating various parts of the circuit from one another, the oxide serves to protect parts of the wafer from dopants during diffusion or ion bombardment (see below).

IA. R. Reinberg, "Dry Processing for Fabrication of VLSI Devices," *VLSI* Electronics: *Microstructure Science,* vol. 2, N. G. Einspruch (cd.) (New York: Academic Press, 1981), p, 1.

In preparation for lithographic pattern formation (illustrated earlier in fig, 8), the wafer is next coated with a thin layer of photoresist—a material analogous to a photographic emulsion which changes chemically when exposed to light (or occasionally, beams of electrons, X-rays, or ions—the latter mostly in laboratory stages of development, as indicated below). The masking and lithographic steps must create the same patterns many times over—figure 3A-1. Masks are typically glass, carrying a grid-like pattern of aluminum or chromium so the mask is transparent to radiation in some areas, opaque in others. SSI and MSI circuits can be made with masks containing hundreds of identical patterns, one for each IC; the entire wafer is then exposed at once. While such a procedure is fast, it cannot produce the narrow lines needed for VLSI; instead,

Figure 3A-1.—Silicon Wafer Showing Patterns Formed by Photolithography

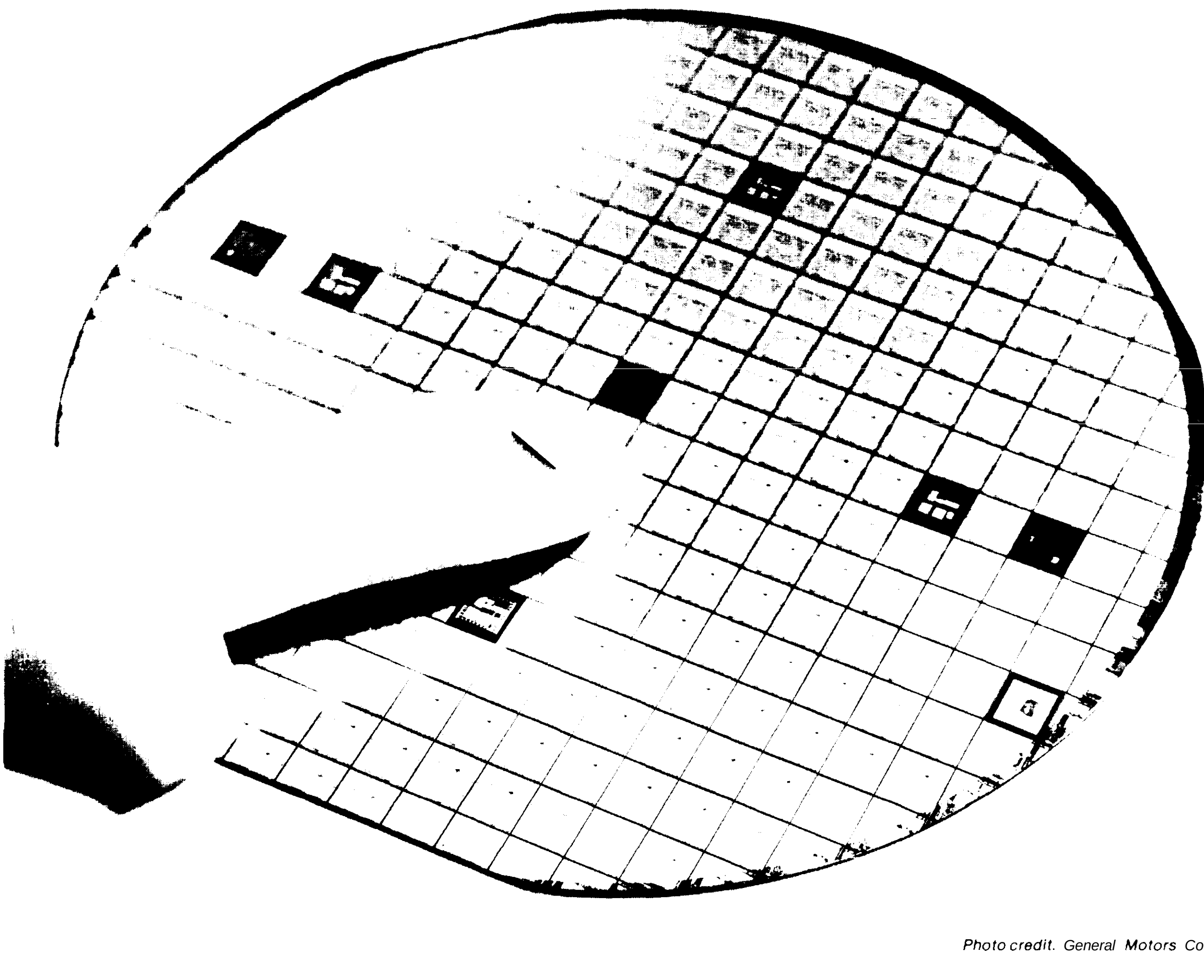


Photo credit. General Motors Corp

the step-and-repeat approach diagramed in figure 8 is commonly adopted. The mask—which maybe 10 times the size of the image on the wafer—creates patterns for only one circuit at a time, after which the wafer is "stepped" beneath the mask and lens. As this process is repeated, a series of identical images in the resist layer is created. The equipment must precisely position and aline the wafer.

Exposure to light or radiation chemically alters the polymeric photoresist. The next step is to dissolve away those portions of the photoresist that have been exposed, creating a pattern identical to that carried by the mask—or, alternatively, to dissolve away those portions that were not exposed. The purpose of creating patterns in the resist is to permit selective etching and doping through the windows that remain. Etching is used to dissolve away material—e.g., portions of the oxide layer—accessible through these windows. In addition to removing oxide, sections of metal layers that have been deposited earlier can be etched away. The metallic layers—fabricated by processes called *metallization—produce* electrically conductive paths to interconnect circuit elements. *Doping* refers to the controlled introduction of foreign elements for altering the conductivity of the silicon. The dopants—e.g., boron, phosphorus—enter the silicon via diffusion or ion bombardment.

During the steps described above, termed wafer fabrication, the entire wafer is processed as a unit. While wafer fabrication can be automated, hand labor is still common; even with automation, the human element is just as important to high yields as the equipment used. When the steps which create the electronic devices and interconnections within each circuit have been completed, the circuits are individually tested—before separation from the wafer—via probes that make input/output connections. Circuits that fail these tests are marked and discarded when the wafer is sawn or broken apart to separate the individual chips. Each good chip is mounted on a chip carrier, which includes pins or prongs for connections to external circuitry, then packaged. An encapsulated IC is shown in figure 3A-2.

By far the largest fraction of defective circuits—hence yield losses—are uncovered at the conclusion of wafer fabrication. These often originate with mechanical flaws such as dust particles settling on the resist, pinholes in the mask, improper spacing of devices so that adjacent circuit elements interfere with one another, or oxide layers that are too thin to provide the necessary insulation, Most such flaws originate in lithography; as the overall size of the chip increases, the probability that it will contain a flaw also rises. This is one of the pressures that leads to greater levels of integration—i.e., making the devices and interconnections on the chip physically smaller, rather than making the chip itself larger, A variety of tradeoffs exist; as chip size increases the yield tends to decrease, but the chip can contain a greater number of functions. When 16K RAMs were coming into mass production, balancing these factors pointed to a chip that was 7 to 8 millimeters square; this gave yields in the range of 20 percent and minimized costs.[2]

Figure 3A-2.—Integrated Circuit Chip Before (top) and After (bottom) Final Encapsulation

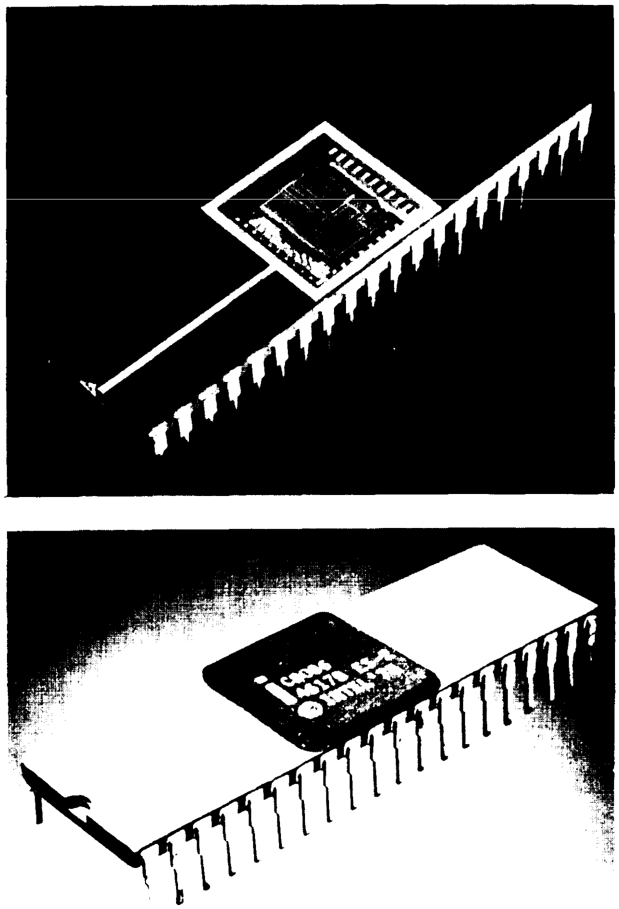


Photo credits: Intel Corp

Beyond the testing that occurs when the wafer is probed before sectioning lie other inspection and testing steps ranging from visual inspection to functional performance checks after the chip has been

[2]J. A. Rajchman, "New Memory Technologies," *Science*, Mar. 18, 1977, p. 1223,

packaged. As ICs become more complex, such tests—many of which are performed under computer control—also become more complicated and expensive. It may be impractical, for instance, to check each memory cell in a large RAM or ROM under conditions that simulate actual operation, much less check all the possible operating states of a VLSI microprocessor. * Algorithms and sampling procedures are used to shorten testing while exercising circuits in realistic fashion. As integration levels increase, designers have greater incentives for building test logic onto the chip to make it self-testing. A parallel trend is to add redundant circuit elements that can be called into play in the event of partial failures.

The lithographic processing described earlier is now done mostly using light. Sometimes critical circuit layers are defined by electron beams, and electron-beam lithography is also used for making masks; in either case, a narrow (micrometer width) beam of electrons "writes" directly on a resist-covered wafer or glass substrate (for mask-making), much as the beam in a cathode-ray tube writes on the phosphor-covered screen. This process is inherently slower than focusing a broad beam of light through a mask to expose the entire chip at once. Moreover, the throughputs—in terms of wafers per hour (table 2)—that can be achieved with electron beam systems are limited by the characteristics of the available resists. Chemical resists that are sensitive to exposure by electron beams require lengthy exposure times, limiting the speed at which the electron beam can write. Electron beams are not the only alternative to light. Among the other candidates are X-rays and ion beams.[3] X-rays have potential advantages in throughput compared to electrons because area exposure through a mask is possible. Unfortunately, intense X-ray sources are not widely available; as for electron beams, resists that can be exposed with X-rays are relatively insensitive, and require long exposures unless the X-ray intensity is high. One source is a synchrotron ring, as used for research on the structure of matter—a very expensive piece of equipment.

Although synchrotron-based lithography for fabricating ICs would be even more capital-intensive than electron-beam lithography, the potentials of X-rays have stimulated considerable R&D. X-ray lithography is likely to become a production tool in the future; electron-beam equipment is already available, and X-rays may move out of the laboratory by the late 1980's. Ion beams have received less attention thus far, but might be able to give resolutions—hence line widths—considerably smaller than either electrons or X-rays.[4] R&D in high-resolution lithographic techniques has been a principal target of government-funded programs in other countries—e.g., the VLSI project in Japan—as well as the VHSIC program funded by the U.S. military.

*A 16K RAM can take on 2'h [384] different logical states—an immense number, nearly $10^{5(XX)}$ Fortunately, the only likely interactions between memory (ells involve those ad jacent to one another, practical ways do exist to determine whether turning "on" one cell will affect data stored near it The problem 15 much more complicated for microprocessors or random logic c hips, whic h lack the regular, repetitive structures character istic of memory.

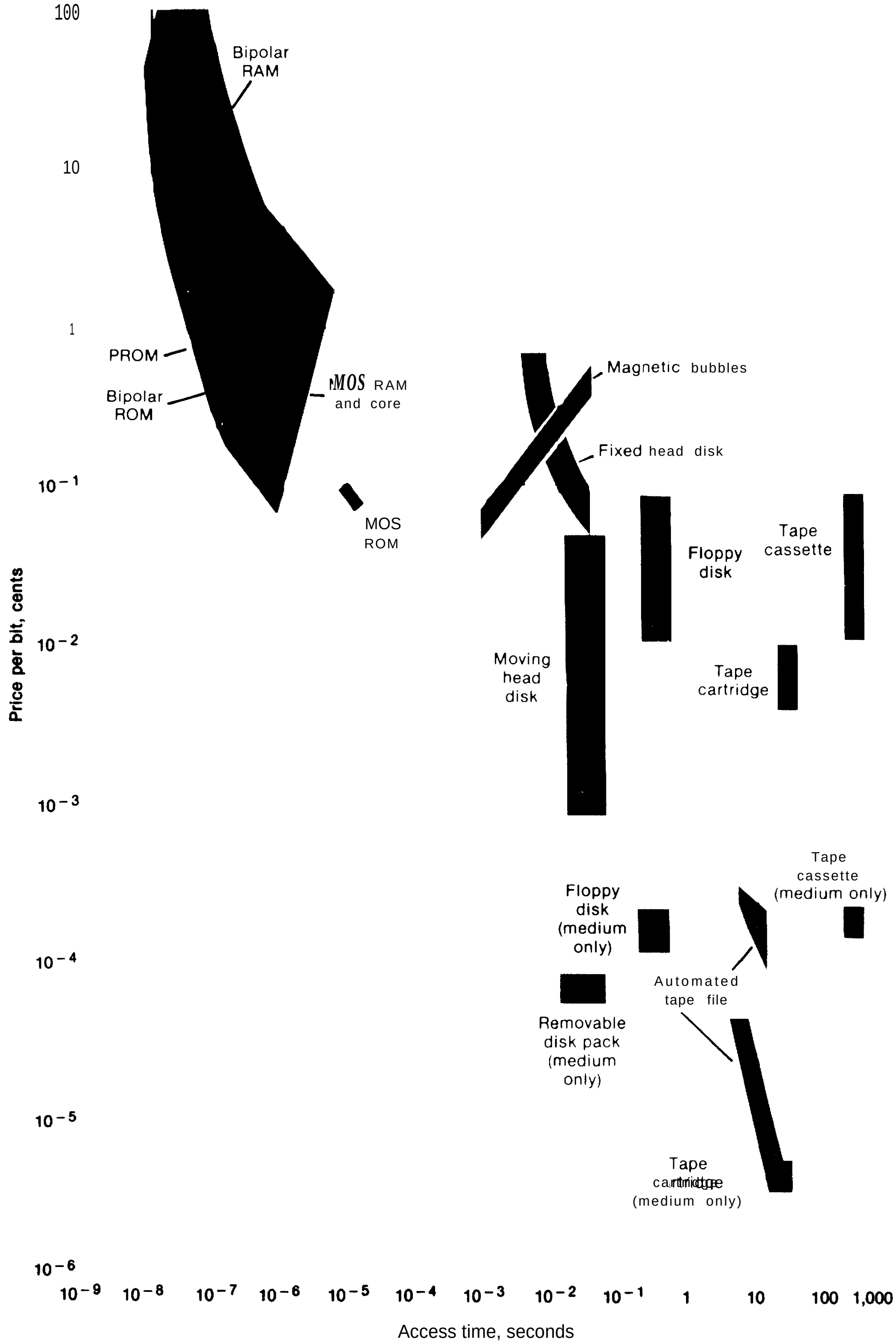
[3] M. P. Lepselter and W T Lync h, "Resolution Limitations in Submicron lithography, " *VLSI Electronics' ,Wi[.restructure Science*, vol 1, op. c it, p. 83, R. K Watts and J R Maldonado, "X-Ray Lithography," *VLSI Electronics. Microstruc ture Science*, VOI 4, N. G Einspruc h (ed) [New York: Academic Press, 1982), p 55

[4] M P. Lepselter, "Submicron Lithography—Limits of Resolution, " *Proceedings, NSF Workshop on Opportunities for Microstructures in Science, Engineering and Technology*. Airlie, ['a., Nov 1 9-22, 1978, p. 187.

Appendix 3B.—Computer Memory

Computer systems make use of different kinds of memory or data storage for different purposes. High-speed cache memory—figure 13 (p. 85)—provided in more powerful computers typically consists of bipolar RAM chips. As figure 3B-1 indicates, these give the fastest practical access times—needed, for instance, for buffering between the main memory and the central processing unit (CPU). Main or primary memory–for storing programs, along with the data being manipulated—generally consists of MOS RAM chips. As figure 3B-1 shows, MOS RAMs are considerably less expensive than bipolar RAMs, though not as fast. A typical IBM model 370/168 mainframe—a large, general-purpose data-processing computer—might have a main memory capacity of 6 megabytes. A byte is equal to 8 bits; thus, the main memory capacity consists of 48 X 10^6 bits—which would require 3,000 16K RAM chips. In contrast, the cache memory for this machine holds 32 kilobytes, or a little over a million bits. The access time for the bipolar cache memory is 80 nanoseconds (80 x

Figure 3B-1.—Performance of Computer Memory Alternatives



SOURCE G. C. Feth, "Memories Smaller, Faster, and Cheaper," *IEEE* Spectrum, June 1976, p 36

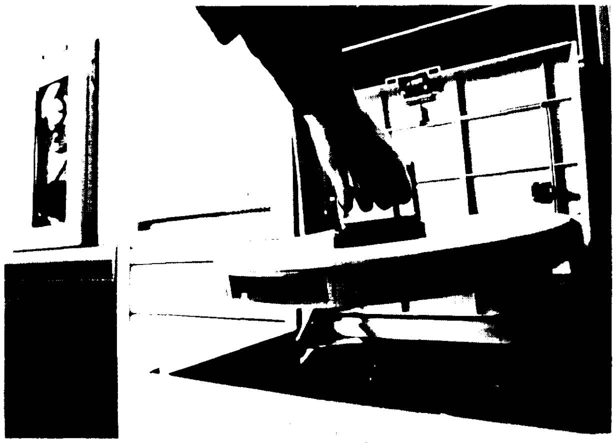


Photo credit Ted Spiege/, 1983

Disk pack for computer data storage

10^{-9} seconds)—corresponding to the cycle time for the 120,000 logic circuits in the CPU—compared to an average of 400 nanoseconds for the MOS main memory.[1] In earlier years, primary memory consisted of magnetic cores (table 3B-1, also fig. 3B-1), wound wire magnets for which the direction of magnetization corresponds to a "O" or a "l." By the mid-1970's memory chips had become cheaper than magnetic cores, largely replacing them except for nonvolatile storage.

IC RAMs are generally supplemented by magnetic disks of various types—table 3B-1. These can store large amounts of information inexpensively; they also provide a form of random access, but are much slower than IC or magnetic core memory (fig, 3B-1).[2] Magnetic tapes are slower yet, but cheaper; large data bases, or records that must be retained for long periods of time, are often stored on tape.

In the future, archival storage may be even cheaper using optical disks. Similar to the Philips video disk, a laser-scanned optical disk about the size of a phonograph record could hold more than 10^{10} bits of information.[3] The 20 million books in the Library of Congress could in principle be stored on 7,000 such optical disks.[4]

[1] W. Anacker, "Computing at 4 Degrees Kelvin," *IEEE Spectrum,* May *1979,* p 26.

[2] See, in general, R. B. J. Warnar, P. J. Calomeris, and S. A. Recicar, *Computer Peripheral Memory System Forecast,* National Bureau of Standards special publication 500-45 [Washington, D. C.: Department of Commerce, April 1979),

[3] K Bulthuis, et al , "Ten Billion Bits on a Disk," *IEEE* Spectrum, August 1979, p 26.

[4] See L. M Branscomb, "Future Computer," *Across the Board,* March 1979, p *61,* who estimates that the hooks in the Library of Congress are equ I\, alent to about 70 X 10^{12} hits,

Table 3B-1.—Computer Memory

Magnetic core: Small wire-threaded toroids are switched between binary states ("O" and "1") by reversing the direction of magnetization. Developed during the 1950's, and now largely replaced by MOS RAMs, magnetic cores were the first inexpensive computer memory that offered fast access times (i.e., microseconds).

Semiconductor: While computer memory could in principle be designed and built using discrete transistors, this would have been much more expensive than magnetic cores. Only when MOS integrated circuit RAMs became available at low cost in the early 1970's did semiconductor memory come into widespread use. By the mid to late 1970's, ICs had become the technology of choice wherever relatively high speeds were called for. Other solid-state storage technologies —e.g., magnetic bubbles—have yet to prove competitive for computer memory.

Magnetic tape: First used with the Univac I in 1951, tape memories are relatively slow but provide inexpensive storage for large amounts of data. The tape is read or written by a recording head much as in an analog audio or video recorder; 1/2 -inch tape drives recording on either seven or nine tracks are common.

Magnetic disk: Digital data can be stored magnetically on rigid or flexible disks. Magnetic drums are also used in specialized applications.

Rigid disks, introduced in 1956, consist of a metal platter coated with a magnetic medium. The disk surface is divided into tracks, which are read and written by a head. In some disk drives, the head is fixed, while in others it moves with respect to the disk surface. Access to particular blocks of data is much faster than for a tape, which must be scanned sequentially.

Removable, or cartridge, disks remain the most common variety of rigid disk, but hermetically sealed "Winchester" drives have also been widely accepted. These can store more data on a disk of given diameter, and are relatively inexpensive, but the disk media cannot be removed for archival storage.

Flexible or floppy disks are made from mylar (a plastic). In function they are similar to rigid disks, but have much lower storage density; they are inexpensive as well as easy to handle and store

SOURCE Office of Technology Assessment

Appendix 3C.—A Process Control Example

Many of the illustrations of computer applications given earlier in table 7 involve multiple processors, either networked or in distributed processing systems. Typical examples include electronic mail or a word processing system with a minicomputer supporting several work stations, Multiple computers are also common in process control. Figure 3C-1 diagrams a portion of the control system for a chemical plant that converts petroleum feedstocks into products like ethylene and butadiene—the latter, in turn, feedstocks for making plastics. Sensors measure parameters such as temperature, pressure, and chemical composition on a continuous basis and send electrical signals to the process control computer. This computer employs a process model and control algorithm—the latter a program that compares the sensor outputs to target values and calculates appropriate adjustments—to monitor and regulate the process in real time. As such, it is a typical example of a feedback control system built around programmable logic rather than hard-wired controllers, Future process control systems will incorporate distributed logic to a greater extent, including smart sensors and actuators, and local controllers linked through networks.

Included in figure 3C-1 is an orifice for measuring flow rate, along with a flow-control valve. The transmitter sends a signal—in this case a voltage proportional to pressure drop across the orifice plate—to the control instrument, which converts it to a flow rate (e. g., pounds per minute). To do so, the control instrument logic must include the relationship between voltage and pressure drop, and the relationship between pressure drop and flow rate—the latter depending on the characteristics of the orifice, In the past, a control loop of this type would normally have relied on analog technology—perhaps even manual readings and adjustments. Now, hard-wired analog systems can be replaced by microprocessor-based digital controllers which are not only much more flexible but also more precise.

The controller in the figure could have a display -e.g., a panel meter—for the plant operators to read, but its primary function is to transmit flow rate data on a continuous basis to the process control computer. This computer monitors many such instruments that read temperatures, pressures, chemical compositions, liquid levels, and other process parameters and adjusts the process according to the control algorithm. In this example, if the flow rate was too high, the computer would return a signal commanding a lower value. The control instrument would convert this to a message that would close down the valve.

In addition to sensors, a typical process control system also includes a number of interlocks analogous to safety valves. If the pressure or liquid level in a reaction vessel or distillation column exceeded preset limits—indicating that the process was out of control, and that the computer had been unable to return the system to the desired condition—the system would automatically shut down so that the plant operators could diagnose the problem, Feedback loops and interlocks improve the control over the composition of the product, as well as helping to maximize yields, efficiency, and safety of operations. Redundancy, reliability, and a system design that allows the plant operators to take manual control quickly and directly in the event of system failures are critical requirements.

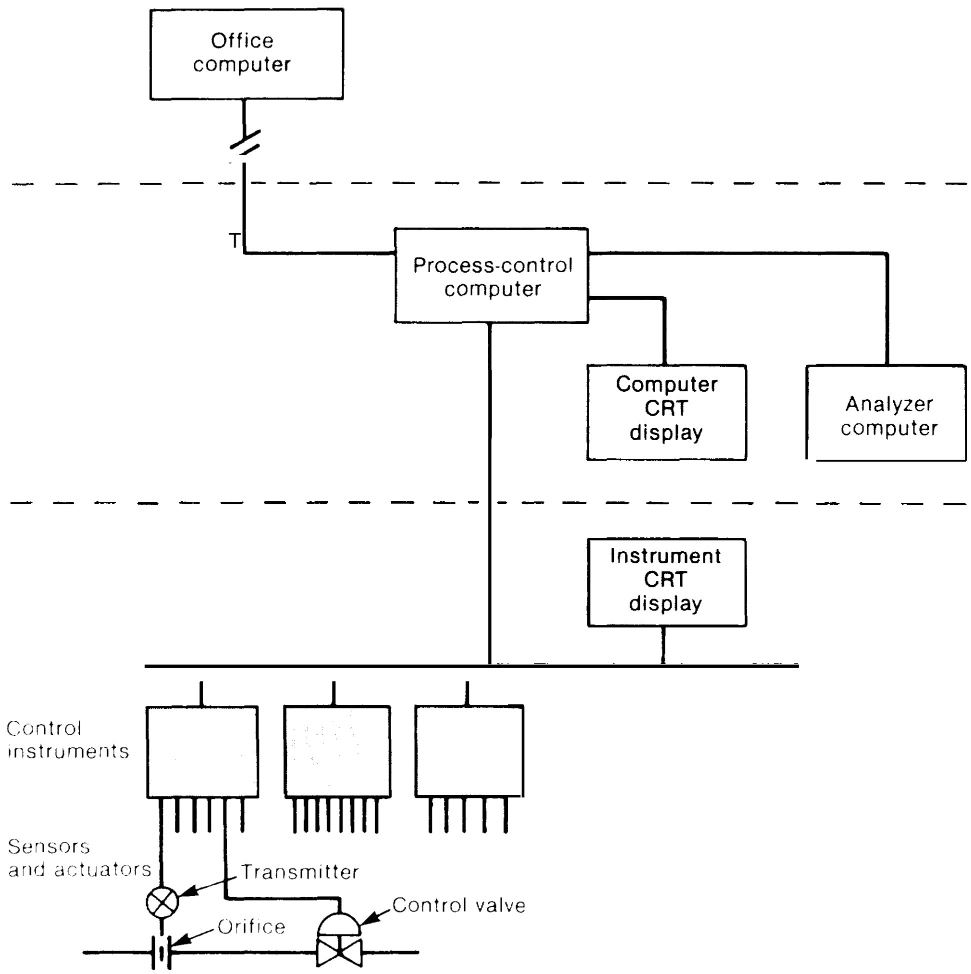
The process-control computer—still in figure 3C-1—can also gather, analyze, and report information on a continuing basis, transmitting it, for example, to the office computer. Once stored, it becomes available for purposes such as recordkeeping or analysis by the engineering department, along with information from the group of other process-control computers spread over a large chemical plant. The instrument CRT in the figure permits a remote operator to monitor a group of control instruments, while the computer CRT is the primary 1/0 device for the plant technicians. The analyzer computer provides off-line modeling and analysis of the process as necessary.

Despite the considerable sophistication of systems like that shown in figure 3C-1, and the further advances expected in the future, some engineers are less than optimistic about the prospects for fully automated plants and factories–where the human operators would be out of the control loop except in emergencies, their usual responsibilities limited to oversight and maintenance. Computerized process control is a field where progress has seldom lived up to expectations; reliability has been a particular problem, and the software available has seldom permitted systems to perform up to the levels promised by hardware developments.[1] The

[1] J J Casso, "Developing a Successful Process Computer System" *Advances in Computer Technology-- 1980*, vol 2 (New York, American Society of Mechanical Engineers, 1980), p 109

Three Mile Island nuclear powerplant accident provides a good example, Total reliance on a computerized system for controlling a nuclear (or nonnuclear) powerplant was then—and is still—impossible. Processes that go out of control are *by* definition system failures—if the process algorithm were adequate, and the system sufficiently robust to withstand equipment failures, control would not be lost. When system control is lost, human operators must intervene,

Figure 3C-1.— Portion **of Process Control System for a Chemical Plant**



SOURCE: Adapted from R. Weber and W. F. Floyd, "Processing Plant: A Hierarchy of Computers and Instrumentation Controls Petrochemical Production," *IEEE Spectrum*, October 1981, p. 56