Electronics and Computers: An Overview

Lewis M. Branscomb

The computer's spectacular growth and growing pervasiveness have been one of the great surprises of modern times.

1) In 1981 6½ million new U.S. cars were fitted with microprocessors. General Motors alone has the capacity to build 25,000 microprocessors a day, making it the largest volume manufacturer of "computers" in the world.

Summary. Microelectronics has been the key to the computer's improved performance and decreased costs. In electronics and computers, "small is beautiful" because small is cheaper, faster, and more reliable. As computers become a daily part of more and more people's lives, emphasis is shifting to programming, ease of use, and applications. Responsiveness to human needs has become the primary challenge facing system designers. How well the industry fulfills this goal will determine whether the emerging information services of the 1980's realize their potential both for improving the quality of life and achieving much-needed productivity gains.

2) An internal message-switching network serving a multinational corporation now links more than 500 computers in over 100 cities in 18 countries, and has been growing at the rate of about one computer a week for the past several years.

3) Computer programmers topped the job demand list in 1981, with a projected 35 percent more jobs available than the year before, according to a survey of employers.

Twenty-seven years ago, when the first electronic business computer was delivered to General Electric in Louisville, it was estimated that perhaps 50 companies in the whole country could eventually use "electronic brains." Today there are more than 1/2 million general-purpose computers in America alone, and available computing power is growing at about 40 percent a year. Three sources of evidence indicate that this rate of growth is likely to continue.

1) First is the increase of distributed processing, driven by the declining cost of computation; the cost of communications, which is also declining but not so rapidly; and the rising cost of any activity involving human labor. Together these trends motivate the use of computing and communications where they save labor, and the use of computer intelligence where it saves communications.

2) The second source of new growth is the use of intelligent systems by the general public. Very large scale integrated (VLSI) circuits (*1*) will be required to increase the level of intelligence in systems for inexperienced users. It takes a great many circuits to make a system easy to learn and friendly to use.

3) The third source is the expanding opportunity for improving office productivity. Systems to help managers and professional information workers have at least 20 times more potential for improving office productivity than today's systems designed to help secretaries. This is partly because there are more principals than secretaries and partly because of salary differentials.

Interdependence of Computers and Microelectronics

The semiconductor electronics and computer industries will become increasingly interdependent as these opportunities are pursued. There can be no technological advance in one without commensurate advances in the other. It is well known that you cannot compete in the broad range of computer products without access to large-scale integrated circuits. Less well known is the extent to which complex VLSI chips of the future cannot be fabricated without automated design, manufacturing, and test systems that are not feasible without computers.

Equally significant is the importance of manipulating, storing, and communicating information in digital form. Digital processing has two advantages over the analog system. First, with digital processing the error rate can be as low as one is willing to pay for in error-correction coding, permitting exact transmission of information. Second, and more important, the digital system permits easy interchange of information. All of the information in computers, data banks, typewriters, and communications networks can, in principle, be translated into digital form and moved from one system to another. In practice, however, useful communication depends on many technical details such as data stream formats and editing conventions that depend not only on the device used to key, print, or display but also on the details of particular applications. Until the applications have been tried and we learn what is economical and what is best for people, it will not be possible to find optimal conventions for useful communications between all these dissimilar information environments.

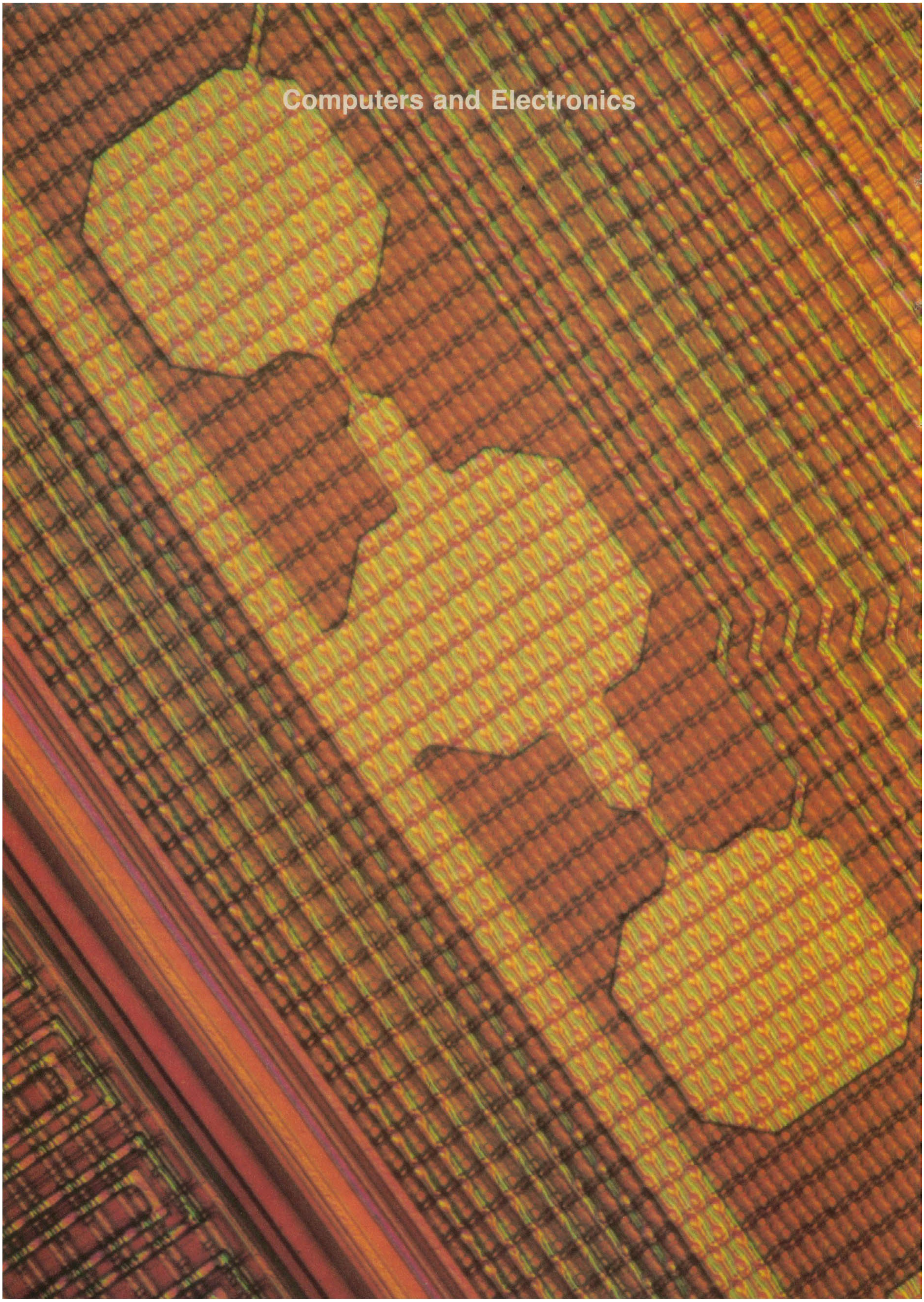
The computer has sometimes been referred to as the productivity engine of our time. The price of small, general-purpose computers of comparable power (Fig. 1), in dollars per instruction executed per second, has been dropping at an annual compound rate of about 25 percent per year since the early 1950's (*2*). The average rate of improvement for the largest general-purpose computers has been about 15 percent per year—less steep than for comparable small machines because the newer large computers offer added function as well as more computing power. The key to this productivity improvement has been microelectronics.

Why "Small is Beautiful"

While the cost of processing a silicon wafer continues to climb with increased wafer size and process complexity, the invention of smaller devices and of lithographic systems for shrinking the smallest feature size have made it possible to build over 1/2 million logic circuits at once on a single wafer. One experimental memory chip (Fig. 2) can store 288,000 bits of information (*3*). This chip stores four times the data, in twice the area, of the 72,000-bit metal gate memo-

The author is vice president and chief scientist of International Business Machines Corporation, Old Orchard Road, Armonk, New York 10504, and chairman of the National Science Board.

0036-8075/82/0212-0755$01.00/0


Computers and Electronics

ry chip employed in the IBM 3600 finance communication system. This 72,000-bit chip is believed to be the largest capacity, dynamic random-access memory component in volume production. Sixty-five of the newer chips, with a total capacity of 18,720,000 bits, have been fabricated simultaneously on a wafer 82 millimeters in diameter.

The real cost advantage of large-scale integration, however, comes from the fact that a computer of a given size requires fewer relatively expensive card, board, and cable interconnections than were needed when each chip contained fewer circuits. Over the years there has been progress in reducing the cost per connection at each level of packaging (Fig. 3), but the greatest progress has been at the chip level where connections are cheapest to start with. By making more connections on the chips, and within the modules that serve to connect the tiny chips to pins that can be plugged into a card, the most expensive levels of packaging can be reduced or eliminated.

Multichip packaging technology dramatically reduces the cost per interconnection by mounting as many as 118 chips on, for example, the single ceramic carrier shown in Fig. 4. With up to 704 circuits on each chip, such a module may contain 45,000 high-speed circuits plus assorted registers and arrays. The large-scale IBM 3081 dyadic system requires only eight of these modules for each of its two processing units, and five each for its controller and channels.

In addition to the speed improvement that accompanies miniaturization, making circuits very small and using dense packaging improves reliability because unreliability comes from the number of connections. That circuits are 10,000 times more reliable today than they were 25 years ago is primarily attributable to higher levels of integration in silicon technology.

Looking ahead, I expect a continued increase in both the number of bits per memory chip and the logic circuit density well into the mid-1980's. Manufacturing costs should continue to drop at 20 to 30 percent a year, although VLSI logic costs may not improve at quite the previous rate because of the complexity of design, testing, and interconnections. This judgment is based on a study of the potential limits (*4*) to continued rapid technical progress. There are four kinds of limits: laws of nature, properties of real materials and processes, technological complexity, and economic realities.

The first are physical limits such as thermodynamic sources, noise, and the

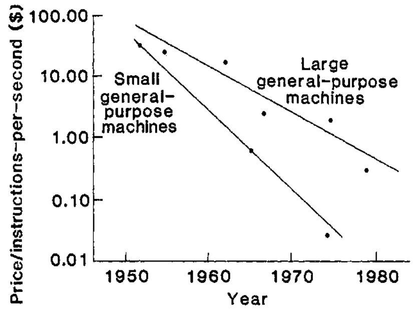


Fig. 1. Trends in computer prices and performance. The lower curve represents an average improvement for roughly equivalent small machines of 25 percent per year; the upper curve shows an average 15 percent per year improvement for large general-purpose machines, newer models of which offer added function as well as more power. Note that the function provided in one "instruction" in the two cases is not equivalent.

speed of light—laws of nature we believe to be immutable. Only the speed of propagation of electrical signals is a serious constraint, and new packaging technologies seem adequate to deal with this.

Technological limits result from the properties of real materials and processes that are related to fabrication cost and yield. A good example is lithography, the pattern-defining process employed in making integrated circuits. The optical lithography tools used in production by the industry today make patterns with minimum dimensions of approximately 2.5 to 3 micrometers (*5*). The practical limit of these optical tools is generally believed to be 1.0 to 1.5 micrometers with the most modern step-and-repeat projection tools. Past this point, hard ultraviolet light, x-rays, or electron beams will be required to fabricate smaller objects with acceptable field size. Materials properties will probably limit useful silicon structures to about 0.25 micrometer in minimum dimensions.

The near-term technical challenge involves complexity limits. How do you design a chip with 25,000 circuits on it? How do you test it in all its possible modes? How would you repair it if it had a fault? Or how can you design it so that it is self-repairing? Managing this complexity requires high-speed computing facilities and challenges mathematicians to invent more efficient algorithms for design, simulation, and testing. Complexity is the main barrier to progress in VLSI.

Equally pressing are economic limits having to do with return on investment. A typical semiconductor production line that cost $1 million for the simpler devices of 1965 might cost $50 million for today's high-density processes (*6*), but of course the modern line has much higher productivity. In 1980, the U.S. semiconductor industry invested $1.4 billion in plant and equipment—representing 17 percent of sales. That is about twice the average investment rate of all American industry. Design costs are up, too (*7*). Some microprocessor chips have cost vendors $20 million to design and develop.

Obviously, in the VLSI world it will be increasingly necessary to manufacture in very large volumes and find ways to decrease design cost and time. To achieve these objectives it will be desirable to select general-purpose designs and to rely on programming for personalizing the systems to specific customer needs.

Importance of Programming

The central electronics complex is actually a small part of an information system—the tip of the pyramid in Fig. 5. The rest of the system hardware, such as disk storage units, tapes, and terminals, is typically substantially more expensive, in total, than the central processor. And none of this hardware will function without the professional programming that controls the system and defines its capabilities. In this area, too, technological progress is being made.

Over the last 10 years, the productivity of systems programming groups has been improving by about 10 percent a year. New programming technology and the use of interactive computer tools for the design and simulation of programming has made possible even faster progress in reducing the number of logical errors that are subsequently found in large programming systems.

Systems programming has traditionally been written in a machine or assembly language for utmost efficiency of operation. This was fine as long as computer architectures based on the discrete electronic components of 15 years ago evolved in orderly fashion. Much of the earlier programming could continue to be used. Today, however, with integrated circuits and microprocessors collapsing the boundaries of a system onto a chip, new system structures and architectures pose a dilemma: one can either take advantage of the new technology and throw out the old software, or constrain the technology by making it run the old code. The answer to this Hobson's choice is provided by optimizing compiler technology.

If the system programming is written

in a high-level language, it can be compiled to machine instructions that are only slightly less efficient than ones coded manually. Now, when new electronics technology dictates a change in architecture, it is only necessary to change the piece of the compiler that generates code for the target machine and recompile.

This still leaves the base of the pyramid in Fig. 5 to be confronted. The application programming and the end user's operation are, in many ways, the largest challenge of the system.

Computer users, over the years, have found that an increasing fraction of their total cost (Fig. 6) lies in programming. This includes both the programming supplied with the computer and the programming done by users themselves, as they evolve from the batch processing of the past to telecommunications-oriented systems, and finally to distributed, complex systems.

It has been speculated that as much as $100 billion has been spent in the field of programming during the past 30 years—about the same as the cost of all the computers installed around the world. Sometime in the mid-1980's, the cost of renting 1 million instructions per second of computing capability is expected to fall below the cost of hiring a professional programmer for the same period of time. It will then become worthwhile to use 1 million instructions per second of computer capability, if it saves even one programmer.

That is one side of the programming challenge: improving the productivity of professional programmers. We also need to make life easier for end users who are not computer professionals. Ideally, a user should not be conscious of "writing" a program at all when asking a computer for information. Providing products that are "friendlier" and more responsive to human needs is a large part of the toughest requirement facing information systems designers today.

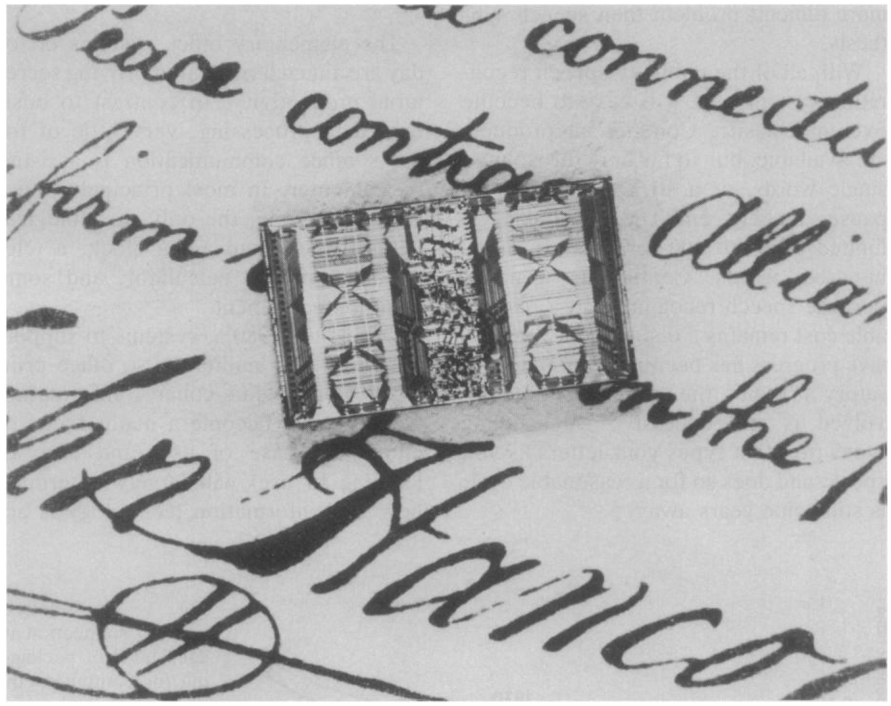


Fig. 2. An experimental computer memory chip capable of storing 288,000 bits of information. The chip is shown against a facsimile of John Hancock's signature on the Declaration of Independence and has enough capacity to store four copies of the Declaration.

Needed: Friendlier Systems

End-user satisfaction is rapidly becoming a decisive competitive factor. That is why the information industry is paying increasing attention to a field of applied psychology called human factors science and engineering, or ergonomics. The main hurdle facing the human factors community is a methodology for measuring and quantifying ease of use. This young science has come far since its first period of rapid applications growth during World War II. The state of knowledge is reasonably good in anthropometrics and the operator productivity of different keyboard layouts, displays, and the like. But there is more to good human factors than productivity. Now, such matters as visual fatigue, noise, and heat are being treated as serious issues by governments, workers, and manufacturers alike. Yet few predictive behavioral principles are established. Tools and measures for quantification of ergonomic qualities are lacking.

The output device that is changing the person-machine interface more than any other, and thus drawing the most attention to ergonomic factors, is the video display terminal. An estimated 5 million to 10 million of them are used by workers in the United States alone. Cathode-ray tubes can, of course, display graphic figures and images in color, and their use has stimulated a demand for all-points-addressable printers that match the capabilities of today's displays. So far, no single technology exists that can meet the entire spectrum of printing needs.

The matrix printer, in which vertical arrays of dot elements move horizontally to produce characters, dominates terminal printer technology. A variety of technologies that do not involve mechanically striking an inked ribbon are also available or in development. Some, such as the electroerosion process in which characters are burned through an aluminum layer to expose a black layer underneath, require special paper. Another nonimpact technology is laser electrophotography—a sophisticated version of xerography.

Inherently simpler, and perhaps most versatile of all, is ink jet printing. Here, tiny droplets of ink are sprayed onto paper, at a speed of up to 100,000 droplets each second, to form characters as sharp as any typewritten ones. Prototype ink jet printers have been built having as many as 256 tiny nozzles, etched in silicon by techniques similar to those used in making microelectronic circuits. The advantage of multinozzle ink jet technology is printing speed, which is roughly proportional to the number of ink jet nozzles times the droplet rate of each nozzle.

Information is usually entered into a computer by means of a keyboard. However, keyboards are not a natural and easy way for many people to communicate. It would be much easier for the general public, for example, if handwritten characters could be recognized by computer. By the mid-1980's recognition of optically scanned hand-printed information should be possible in a range of applications.

The most natural way for people to deal with information systems may be through speech. Computers today can be made to talk quite well, but, like many of us, machines still find it hard to listen. Speech recognition is clearly a more difficult problem than speech synthesis.

With all of the publicity speech recognition systems get, it is easy to become overenthusiastic. Commercial products are available, but so far only for isolated single words, or a string of them with pauses in between. The vocabulary is limited to about 100 (or at most a few hundred) words. Continuous, general-purpose speech recognition at a reasonable cost remains a distant goal. Impressive progress has been made in the laboratory (*8*), but the computer time involved is still excessive. A listening typewriter that types your letters as you speak, and does so for a reasonable cost, is still some years away.

Office of the Future

The elementary office systems of today are aimed largely at improving secretarial productivity. In contrast to business data processing, very little of today's office communication is assisted by computer. In most principals' offices, for example, the only technological items to be found are a clock, a telephone, a pocket calculator, and some dictation equipment.

During the 1980's, systems to support the world's 30 million or so office principals, or "white-collar" information workers, will become a major focus of effort (*9*). Ease of use (and ease of learning to use) will largely determine how well information technology is accepted by managers and professionals.

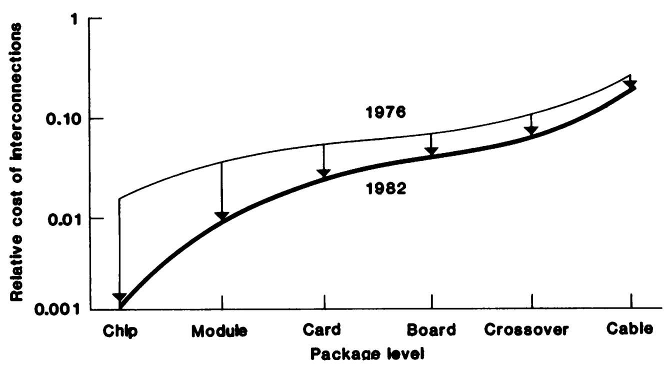


Fig. 3. The relative cost per connection at each level of packaging for computer circuits. In typical packaging, a chip is mounted on a module; modules plug into boards, which plug into cards, which are connected to other cards through crossovers and cables. Progress has been made in reducing costs at all levels, but especially at the chip level through microfabrication. The more connections are made on chips and modules, the fewer of the relatively expensive connections a computer of a given size will require.

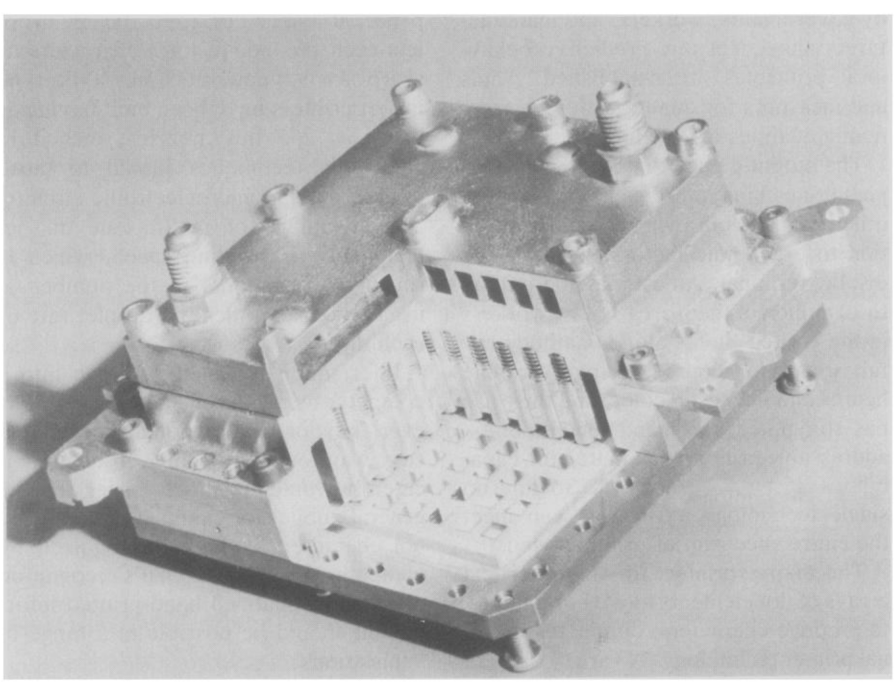


Fig. 4. A thermal conduction module containing 118 computer logic chips, 20 of which are visible in the cutaway section. Within the module, spring-loaded pistons in contact with each chip convey heat through a metal plate to a circulating chilled water system. Modules like this plug into boards which permit more than 3/4 million logic circuits to be contained in about 4 cubic feet.

Eventually, office systems will have to be able to merge the separately handled data, text, voice, and image information into integrated electronic documents that can be communicated, retrieved, and otherwise dealt with without the user's being conscious of its form when stored. Artificial distinctions between data, word, and image processing will gradually disappear as document distribution systems—for all kinds of documents—are evolved.

Another aspect of improving office communications is human factors at the systems level. An example is the matching of information network structures to the needs of real organizations. In the early days of data processing the computer sat behind glass walls and was operated in a batch, off-line mode by a central authority. The next step was a time-shared environment in which the central computer served all units of the enterprise over teleprocessing communications lines but was mapped to the formal authority structure of the company's organization chart.

A modern computer network should be able to reflect equally well the many informal peer relationships that bypass the authority structure, as well as the structure of formal transaction applications. Today's flexible system network architectures (*10*) allow the information flow to adapt to real organizations and the needs of their people, instead of making the organization conform to the system's structure. Networks can vary from loosely peer-coupled arrangements where work flows through the network without any node being in control, to sophisticated, top-down networks with tight central control. No one approach is best for every organization, and different application structures should be able to coexist in the same basic transmission facilities.

Communications for the 1980's

During the 1980's, work on communications system architecture must focus on three fronts: communications within an establishment, the geographically distributed communications network outside, and the gateways between these two. Ideally, to a user, the distinction between a local message destination and a remote one should be invisible.

One technique for local area communications that allows many devices to share a digital transmission link is the contention bus, typified by Xerox's Ethernet. Here, a device tries to transmit; if another device is sending, the first device waits and tries again. Another approach is called the token ring. Here, the device holding the "token" has the right to transmit; as soon as it is finished, the token passes to another device. A third technique, buffer insertion, works by having any device with something to say temporarily buffer all incoming traffic while it sends its message. Still another approach is the broadband cable television, frequency-multiplexed facility. Sorting out how the communication within an establishment should be handled will be a major issue of the 1980's (*11*).

The second area of interest is the character and function of system gateways—nodal connections that transfer information back and forth between a local area network and the larger network outside. The relationship of these gateways to programmable, digital branch exchanges is another source of debate.

The third area on which work must focus is the network service connecting dispersed establishments, including both private networks and the public digital communications services and data networks now offered or planned in many countries. Today most data traffic is still carried on facilities designed and optimized for voice telephone service. The new digital services offer improved accuracy and richer function and eliminate the need for modems, or analog-to-digital converters.

Two broadband communications technologies will become important in this decade. Communications satellites offer a channel cost that is falling at 40 percent a year as demand rises. The information inside a computer runs at millions of characters per second, but can be accessed at only a few thousand characters a second over telephone lines. The satellite allows computers to communicate at the same speed at which they operate. Satellites are also making it possible to combine data-processing applications with digital voice communications, office applications, such as facsimile, and video applications, such as teleconferencing.

Eventually, buried, bundled optical fibers will provide another attractive alternative for long-haul transmission systems spanning continents and oceans with vast amounts of bandwidth, error-free digital transmission and improved security. Optical fibers will find much earlier application, however, for high-speed channel communications within an establishment and perhaps as a backbone for local area networks.

Emerging Information Services

During the 1980's, satellites and computers will join with cable television systems and telephone lines in delivering low-cost "videotex" information services directly into businesses and homes. The terminals will be either ordinary television sets with adapters and key pads, or displays designed for the purpose. Some of these systems, like Canada's Telidon, will be interactive. Initially, these systems are likely to be used for simple business applications by those who do not require the function, complexity, and cost of their own data- or word-processing systems. This will be possible because videotex systems can include applications accessible only to a group of subscribers.

Another new medium that may provide inexpensive access to large libraries of stored information is the laser video disk. Pressed in plastic, a disk can contain on each surface 54,000 color television images—equivalent to 10 billion bits of coded information (*12*)—plus two channels of stereo sound. The contents of a 300-page book could, in principle, be encoded in each square centimeter of surface.

Obviously, videotex and the video disk both depend upon the generation and availability of suitable program material, or software. As the distinction between computing and communications services diminishes, so too will the distinction between information services and information-handling services. The 1980's will no doubt find more suppliers of information handling equipment both in business with information vendors and in the information business themselves.

The largest number of information machines, like the dozens of electric motors scattered throughout a typical household, will be totally invisible. At some point in the future, many computers will simply disappear from view, having been dispersed into the appliances around us. The computer will vanish into the typewriter, into the telephone, into the thermostat and toaster. That does not mean I will have to write a program to run my "smart" toaster; it just is not going to burn the toast any more.

Thus, the computer is becoming something much more than a productivity engine. It is on the way to becoming a responsive personal servant to ordinary people, at home as well as at work.

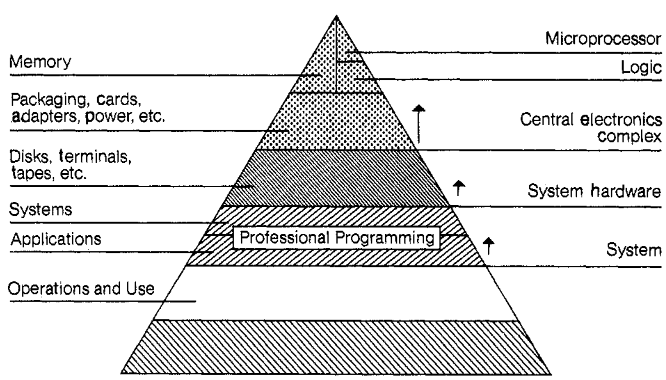


Fig. 5. The system pyramid. The central electronics complex represented at the top is only a small part of a modern information system. Other system hardware is typically more expensive. Professional programming is also required to control the system and define its capabilities. The large base of the pyramid represents the application programming which computer users must undertake and the terminals and other distributed processing equipment that are bringing computing capability to increasing numbers of people.

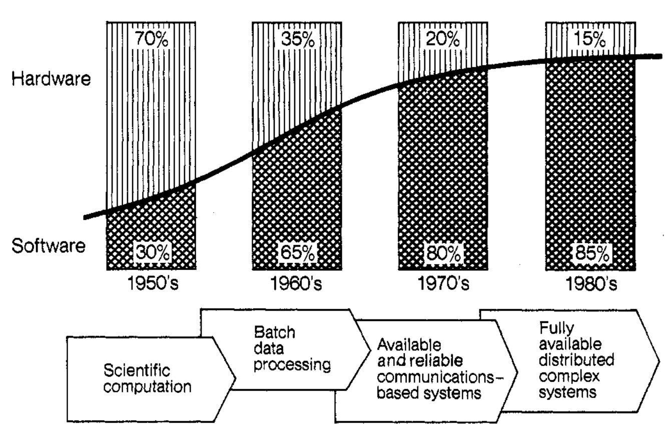


Fig. 6. Trends in computer hardware and programming costs. In the long run, programming accounts for more of a typical computer user's cost than the hardware. This programming includes that built into the computer as well as that done by users themselves as they evolve from batch processing to telecommunications-oriented systems and, finally, to distributed, complex systems.

Microelectronics is part of the answer to making information systems both affordable and sophisticated enough for widespread use, but the innovation challenge is not just technical. It is matching the broad spectrum of new technologies to opportunities for serving the information and education needs of contemporary society.

References and Notes

1. Definitions of VLSI abound, but *very* large scale integration is sometimes considered to start at about 5000 MOSFET (metal oxide semiconductor field-effect transistor) logic circuits, or more than 16,000 memory bits, per chip.
2. P. F. Olsen and R. J. Orrange, *IBM J. Res. Dev.* **25** (No. 5), 405 (1981).
3. Paper to be presented at the International Solid State Circuits Conference, in February 1982.
4. V. L. Rideout, *IEEE COMPCON 80 Technical Digest* (Institute of Electrical and Electronic Engineers, New York, February 1980), pp. 2–5.
5. ——, in *Digital Technology Status and Trends*, H. Pinke, Ed. (Oldenbourg, Munich, 1981), pp. 11–48.
6. A. L. Robinson, *Science* **208**, 480 (1980).
7. P. M. Russo, *IEEE Trans. Electron Devices* **ED-27**, 1332 (1980).
8. F. Jelinek, *Proc. IEEE* **64**, 532 (April 1976).
9. P. A. Strassmann, *Technol. Rev.* **82**, 54 (December/January 1980).
10. L. M. Branscomb, *Comput. Networks* **5**, 3 (1981).
11. N. Mokhoff, *IEEE Spectrum* **18**, 57 (April 1981).
12. L. M. Branscomb, *J. Commun.* **31**, 143 (winter 1981).

Computers: A Survey of Trends and Limitations

Joel S. Birnbaum

Summary. Enormous progress in electronic technology is accelerating the use of computers in everyday life. In this article trends in hardware, input-output technology, computer architecture, software, communications, and artificial intelligence are examined and complexity is identified as a limitation to further progress. Promising directions of research, which may extend the range of computer applications, are discussed.

The popular press and some computer scientists would have us believe that we are on the verge of the realization of an age-old dream: intelligent, obedient machines relieving people of much of the hazard and drudgery of life and extending mental capabilities and creativity through the filtered delivery of information from a global range of sources. The evidence presented for this is the growing use of inexpensive microprocessors in homes, laboratories, and factories, the availability of new communications technologies, and the extrapolation of the everyday uses of computers as calculating engines for scientists and as data processors for businessmen. Increasing computer literacy in the current generation of students is expected to multiply the rate of development of new applications. Yet many of the limitations to our progress today are precisely those that have prevented many of the sanguine projections of a decade ago from coming to pass. These are not fundamental limitations imposed by physical laws; rather, they result from our difficulty in dealing with complexity. Software development productivity has suffered most from human-introduced complications (see Table 1), compounded by approaches that are seldom disciplined or scientific (*1*).

In this article trends in computer technology, including hardware, interfaces, architecture and organization, software, communications, and artificial intelligence are examined, and major changes that seem likely in the design and application of computers are indicated. The improvements in both hardware and software technologies are now so great that basic reformulation of our design assumptions may be necessary, and greatly increased attention to tools and technology to combat complexity is called for.

Trends in Hardware: Logic, Memory, and Storage

The startling advances that have been made in logic, memory, and storage technologies are all possible because a binary digit has informational value independent of its physical size. Technologists have produced ever faster, cheaper, and less power-hungry logic and memory elements simply by shrinking them; much of the research in computer technology can be thought of as a quest for smallness. As objects change radically in size, scaling relations between surface area and volume become important, and much new science has been driven by a need to understand the nature of surfaces and thin films and the nature of the interactions at interfaces between dissimilar materials.

As critical dimensions approach 1 micrometer, the wavelength of light becomes a limiting factor to attainable lithographic resolution. Progress in electron beam– and x-ray–based techniques will circumvent this problem, and most technologists do not think lithography will prove to be a limitation for the next 10 years (*2*).

A more serious limitation is signal propagation speed, which is about 1.5 nanoseconds per foot in contemporary packages. To achieve 1 billion instructions per second in a conventionally organized machine, the basic instruction cycle of 1 nanosecond can have a maximum path length of only a few inches. Very fast conventional machines must therefore be extremely small, and so component interconnection assumes paramount importance. Densely packed devices running at high speeds require efficient cooling mechanisms. One of the boldest attempts to solve this problem is being pursued by researchers, primarily at IBM, who are investigating superconducting Josephson junctions that promise two to three orders of magnitude less power consumption than conventional silicon technologies. While the laboratory demonstrations of this technology are very impressive, the difficulties imposed by extraordinarily critical manufacturing tolerances and the need to

The author is director of the Computer Research Center, Hewlett-Packard Laboratories, Palo Alto, California 94304.

0036-8075/82/0212-0760$01.00/0