VOLUME 5 □ NUMBER 4 APRIL 1968

IEEE spectrum

features

Departments: *please turn to the next page*

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.

departments

the cover

Computer-driven photocomposition machines now permit the generation of type in any face and font at electronic speeds. The artwork attempts to depict the versatility now possible in this new era of the graphic arts. For the story of the quiet revolution in the printing and publishing industry, turn to the first installment, which begins on page 48.

Rosebud, clover leaf, saguaro, fir-tree, butterfly, batwing, signpost, turnstile—these and other poetic names have been given by antenna engineers to their creations. While the names of many antennas are commonplace and need no description, the more exotic varieties, whether figurative or literal, could advantageously have a glossary to refresh the memory. Antenna jargon has the merit of being lighthearted. Perhaps the spirit of the antenna creator can be captured as we look at a few entries which might be put in an antenna glossary. The following list compiled by Jack Ramsay, AIL's consultant in our Department of Radar Techniques and Advanced Development, is a miscellany representing only an interesting fraction of the total variety of antennas.

The Variety of Antennas

Anntennafier: Antenna plus amplifier combined in single package.

Antennamixer: Antenna incorporating frequency converter.

Antennaverter: Same as *Antennamixer*.

Barrel: Microwave cosecant theta reflector, of which one half is a paraboloid, and the other half is formed by rotating the principal parabolic section through 90 degrees about a line through the focus perpendicular to the axis.

Batwing: Slot antenna with flanges having b. shape to control radiation pattern.

Beavertail: Nodding height-finder antenna having b. beam shape.

Bed of Nails: Surface-wave antenna of pins on a groundplane.

Beercan: Mast or antenna constructed of b.'s soldered together.

Billboard: Planar array of dipoles with reflecting screen. (cf. *Curtain, Fir-tree, Mattress, Pine-tree*)

Birdcage: Polarized fishbowl shaped reflector constructed of parallel wires or slats at 45 degrees to the equator, and conforming to the curved surface. A feed within reflector polarized at 45 degrees illuminates a portion of the b., and produces a reflected, collimated beam which passes through the opposite side of the b. since this side is orthogonally polarized. Appropriate rotations of feed only give 360 degree azimuth scan and some elevation scan. In various versions also called "polarized barrel," "parabolic dome," and *Helisphere*.

Blinders: Shields at sides of antenna aperture. (cf. *Tunnel*)

Bootlace: Lens antenna formed by back-to-back arrays linked by transmission lines and/or phase shifters.

Bow Tie: Dipole or slot antenna formed by two triangular areas in b. configuration.

Butterfly: Two *Scimitar* antennas joined at their bases and feed points, and making equal angles with a groundplane.

Cage: Thick dipole (or monopole) of uniform or non-uniform section constructed of wires held by horizontal rings.

Cartwheel: Top capacitance of dipole (or monopole) formed by radial spokes terminated by circular ring.

Cheese: Short parabolic cylinder between parallel plates, the polarization being parallel to the plates. (cf. *Pillbox*)

Cigar: End-fire array of disks of varying diameter giving a c. shape profile.

Clam-shell: A geodesic Luneburg lens resembling the *Tin hat* but having part of the brim facing inwards.

Cloverleaf: Four adjacent loop antennas in a plane.

Cobra: Center fed dipole fed by coaxial cable through lower half. Cable is coiled at base of antenna to form decoupling choke.

Cornucopia: Long horn feeds offset reflector at 45 degrees, two extended sides of horn providing shields. (cf. *Hoghorn*)

Curtain: Vertical array of driven or reflector elements suspended from triatic.

Discage: Composite wideband structure combining a *Discone* antenna on top of a *Cage* antenna of *Cobra* fed type.

Discone: Broadband dipole consisting of a conical lower *Skirt* and an upper short vertical section loaded with a capacitive disc.

Dish: Concave reflector antenna with circular periphery.

Ditch: Flush antenna obtained by exciting d. in ground, or groundplane.

Dumb-bell: Resonant slot radiator having rectangular central portion and enlarged circular ends. Also applied to tubular waveguide having d. cross section.

Fakir's Bed: See *Bed of Nails*.

Ferrod: Tapered rod of microwave ferromagnetic material fed by waveguide.

Fir-tree: Vertical array of horizontal dipoles fed by transposed two-wire line and backed by reflector array. (Ger. Tannenbaum)

Fish-bone: End-fire array of non-resonant wires lightly capacitively coupled to two-wire line.

Fish-eye: Spherical lens of variable index fed at one end of a diameter and focusing at opposite end.

Fly-swatter: Upper reflector, planar or curved, at top of mast of *Periscope* antenna.

Funnel: Original name of electromagnetic horn (1897).

Gable: Antenna illumination function of isosceles triangle shape.

Half-cheese: Offset fed parabolic cylinder between parallel plates. (cf. *Cheese*)

Halo: Folded dipole bent into circle with ends capacitively loaded.

Handle: A rectangular half loop over a ground plane.

Helisphere: 360 degree scanning microwave reflector consisting of a spherical radome with closely spaced helical wires of 45 degree pitch angle embedded in surface. (cf. *Bird-cage*)

Helmet: A geodesic Luneburg lens resembling the *Tin hat* but having part of the brim facing downwards.

Hoghorn: Sectoral horn integrated with half-cheese has hoglike shape. Applied also to *Cornucopia*.

Holey Plate: Metal plate perforated uniformly or non-uniformly with holes, and used in linear arrays, surface wave antennas, array simulations and quasi-optical filters.

Hourglass: Reflector formed by rotating parabolic curve around an axis behind the vertex and parallel to the directrix.

Hulahoop: Short vertical antenna top loaded by horizontal wire bent into a nearly complete circle, the end being connected to ground by a small tunable capacitance.

Hurdle: A low height, vertically polarized array of resonant rectangular half loops over ground or a groundplane.

Indian Club: Broadband monopole of bulbous shape.

Janus: Antenna having independent beams pointing fore and aft from a mobile craft.

Ladder: Resonant array of wires in L structure.

Lance: Microwave rhombic antenna used as array element.

Lazy-H: Two horizontal full wave dipoles stacked in elevation with vertical feed line.

Letter-rack: Cavity antennas stacked in E plane and series fed by transposed loop couplings from transmission line.

Manipole: A broadband omnidirectional antenna consisting of a plurality of vertical elements of different lengths in close proximity.

Mattress: See *Billboard*.

Nitch: A *Notch* antenna having a length less than quarter wavelength.

Notch: A quarter wave resonant open end slot cut in the edge of a conducting surface, as an airplane wing, and transmission line fed.

Parant: Antenna system containing an integral parametric amplifier.

Periscope: Two reflector antenna usually consisting of paraboloid antenna radiating vertically at inclined *Flyswatter* reflector mounted on mast. Horizontal p. antenna also used as radio telescope.

Pillbox: Short parabolic cylinder between parallel plates perpendicular to axis, with polarization perpendicular to plates. (cf. *Cheese* and *Half cheese*)

Pine-tree: See *Fir-tree*.

Pinwall: Horn antenna with two opposite sides constituted by spaced rods with non uniform spacings.

Pyramidal Horn: Horn antenna of pyramidal shape, fed at vertex.

Pylon: Thin vertical cylinder containing single vertical slot, line fed, yielding omnidirectional pattern.

Quad: Resonant square loop with square loop parasitic reflector. Also applied to single loop.

Quadraloop: Four rectangular loops on cylinder in planes containing axis of cylinder.

Quadriped: Antenna with four legs.

Quadripod: Four leg feed support for reflector antenna.

Rabbit's Ears: TV receiving antenna consisting of adjustable V dipole.

Radant: Antenna built into radome, or antenna which is its own radome, as dielectric lens and *Helisphere* antennas.

Rainspout: Inexpensive mast or antenna made from corrugated r. tubing.

Rectenna: Power converting antenna consisting of array of elements each terminated in rectifier, with outputs added powerwise.

Rocket: Omnidirectional slotted cylinder antenna (archaic).

Rosebud: Circular arrangement of inclined *Scimitar* antennas with common narrow end feed points and separately grounded broad ends.

Saguaro: Vertical broadband linear radiator with coplanar inclined branches of different lengths.

Sausage: Leaky helix plus transmission-line antenna resembling s. (Fr. *Saucisson*)

Serpent: Convoluted traveling wave slotted waveguide array with high dispersion.

Scimitar: Half loop antenna of scimitar shape fed at thin end and grounded at broad end, or vice versa.

Shovel: Microwave reflector forming shaped beam.

Signpost: Two slotted waveguide arrays in V configuration.

Skirt: Lower half of half-wave dipole formed by folded back section of outer conductor of coaxial line, inner unshielded section of which forms top half of dipole, also called "coaxial antenna" or *Thunderstick*.

Sleeve: Cylinder enclosing central region of dipole or lower end of monopole antenna.

Snake: See *Serpent*.

Stub: Short antenna extending from groundplane.

Thunderstick: See *Skirt* dipole.

Tin-hat: A geodesic Luneburg lens consisting of two metal surfaces of resolution of t. shape with constant normal spacing. At the brim the surfaces are bent so that the opening faces outwards horizontally.

Top-hat: Cavity radiator with circular flange (1891).

Trough: *Ditch* antenna, also waveguide with one wall removed, and containing central fin with exciters.

Tunnel: Cylinder attached to circular aperture antenna to reduce side radiation (cf. *Blinders*)

Turnstile: Two coplanar dipoles as horizontal X, fed in quadrature to produce rotating field.

Umbrella: Vertical radiator with top capacitance loading by inclined radial wires. Also u. paraboloid is a symmetrical reflector erectable like an umbrella.

Valentine: Two coplanar *Scimitar* type antennas with broad ends joined and narrow ends brought together to form feed point.

Whip: Flexible vertical antenna mounted on vehicle or man-pack radio.

Wire Slot: Resonant rectangular loop, or half loop over ground

Wullenweber: Direction finding circular array around cylindrical reflector. Limited sector of delay line collimated elements can be positioned on any azimuth.

Zeppelin: End fed half-wave dipole with quarter wave matching section.

IEEE spectrum

Spectral lines

Three problems for the profession. Like the IEEE, the American Physical Society customarily sponsors not only a succession of smaller meetings, but also a great annual meeting in New York. Since about a third of the papers submitted to IEEE publications come from authors who have been trained as physicists, there is a presumption that any important meeting of physicists is of interest to a proportion of readers of IEEE SPECTRUM. At the 1968 meeting, which was held in Chicago, overlap with the interests of electrical and electronics engineers was particularly strong.

Along with many topics not—or, at least, not yet!—of interest to engineers (such as "A Quark Model for Nonleptic Hyperon Decays"), three that were clearly of first importance to the physicists are also important to many engineers. They were financial support, recruitment to the profession, and publication.

It was pointed out that in respect to financial support, physicists in the U.S. are to some extent the victims of their own policy. In their relations with the federal government, American physicists have taken the position that in order to be healthy, science requires budgets that grow at a constant rate. It turns out, however, that scientists and engineers are not the only ones who understand the exponential function; its properties are appreciated also by certain people in the Bureau of the Budget. Now that research in science is costing a discernible fraction of the national income, there are questions in Washington about why it has to keep growing at all.

Traditionally, the argument has been that science is useful, and therefore worth supporting. However, Project Hindsight has shown, purportedly, that useful technology lags undirected science by about 40 years. In 40 years, most of those who are now determining policy will be dead—and even if not, they certainly will not be running for re-election—so that era is of little interest to them. Hoyle, the noted astrophysicist and author of science fiction, puts succinctly the fundamental conflict: he says that scientists are motivated by curiosity, whereas society is motivated by greed. He asserts that the more civilized a society is, the more it has retreated from the natural world, and the more hostile it is to science.

The concern for education was triggered by a decline in the numbers of students majoring in physics in the U.S. Your editor was reminded of a pronouncement made some 30 years ago by a fabulously salaried Hollywood executive when interviewed about a giant slump in attendance at the movies: "As soon as the money stopped coming in, I knew that something was wrong." The physicists now sense that something is wrong, because students are coming to them in smaller numbers.

Professors of engineering are having the same trouble; some of the reasons may be different, but some are undoubtedly the same. One reason is the identification of physics and engineering with war-making and destruction. Another consideration (no doubt the other side of the coin) is the widespread desire to take immediate, positive, social action by entering government, joining the Peace Corps, or some similar course. Actually, there are few, if any, modes of undergraduate preparation for the Peace Corps that would be more useful than a major in physics. However, physics in Academe has become so imbued with specialism that a student who does not expect to go for a Ph.D. degree is considered frivolous and not worth bothering with. Consequently, the A.B. in physics has become a preparation for graduate school in physics—and anyone with enough brain to be a physics A.B. has enough brain not to spend four years in preparing for schooling that he does not want. These considerations may have some relevance for the developing engineering curriculum.

One of the unremarked ironies of the meeting was that in one ballroom the leaders in teaching debated how to cope with the defection of students to the socially oriented disciplines or indisciplines, while in another leaders in research advocated that physicists drop physics and help to clean up the sociological mess. "Just imagine," said one, "that ceremonial pitch that opens the baseball season, and consider what would happen to baseball if pitching like that went on all summer."

The third large question addressed by the physicists was the problem of publication. More is being published than anyone can learn, but that has been true for a long time; the question is whether you can find what you need to know. In the fields in which I have worked, the abstracting journals and personal contacts can keep one from missing anything important to one's own effort. I am now convinced, however, that in some areas (and Conyers Herring cited convincingly the area of the solid state) an information problem does exist objectively, outside the minds of librarians and other inventors of paperwork. Even if it didn't, publication on the present scale would raise some questions demanding notice. Who should pay for it? Does a need for it really exist? If so, is journal publication the best way to satisfy the need? Can it be paid for by the persons who need it?

These questions are receiving a great deal of organized attention, in many places. A revolution in the recording and the finding of technical information is on the way. In the IEEE, no problem is receiving more thought, and there has already been a significant amount of action, which this column will soon report. *J. J. G. McCue*

Authors

High-field superconductor technology (page 63)

Charles Laverick is a senior electrical engineer at the Argonne National Laboratory, where he has been a staff member since 1958. He has received international recognition for his contributions to superconducting-magnet technology, which include the development of copper stabilized superconducting cables and the operation of several significant high-field superconducting coils. His early work at Argonne was devoted to the development of prototype low-power-level frequency programming and RF accelerating cavity biasing systems for the 12.5-GeV zero-gradient synchrotron. He graduated in physics from the University of Durham, England, in 1950. From 1950 to 1955 he was a senior physicist and group leader in the Research Laboratories of the General Electric Company in England, where he developed modern environmental testing techniques for airborne radar and missile guidance equipment. He later spent a year with Canadian Industries Ltd.

A prediction of power system development, 1968 to 2030 (page 75)

Alexander Kusko (SM) was born in New York, N.Y., on April 4, 1921. He received the B.S. degree in electrical engineering from Purdue University in 1942 and the S.M. and Sc.D. degrees in electrical engineering from Massachusetts Institute of Technology in 1944 and 1952, respectively. From 1944 to 1946 he was an electronics maintenance officer with the United States Navy. From 1942 to 1944 and from 1946 to the present time he has served on the staff of the Electrical Engineering Department at M.I.T. He has been in charge of the M.I.T. network analyzer, the M.I.T. Electrical Machinery Laboratory, and Air Force-sponsored projects on aircraft machines and regulators. He is also the president of Alexander Kusko, Inc., consulting engineers, Cambridge, Mass., and a lecturer in the Electrical Engineering Department at M.I.T.

Dr. Kusko has worked in the fields of power system analysis, power solid-state and magnetic circuits, compressed-gas insulation, underground power transmission, and rotating machinery, among others. He is the author of a number of technical papers and has been issued 13 U.S. patents. He is a member of Sigma Xi, Tau Beta Pi, and Eta Kappa Nu.

New faces of Eve: women in electrical engineering (page 81)

Irene C. Peden (SM) received the B.S. degree in electrical engineering from the University of Colorado in 1947. She worked for the Delaware Power and Light Company for two years before joining the staff of the Stanford Research Institute, where she worked on antenna problems. As a research assistant at the Stanford Microwave Laboratory from 1958 to 1961 she worked on measurement techniques for microwave periodic circuits. She received the M.S. degree in 1958 and the Ph.D. degree in 1962, both from Stanford University. In 1961 she joined the faculty of the University of Washington, where she is now an associate professor of electrical engineering.

Dr. Peden is a senior member of the Society of Women Engineers and a member of Sigma Xi. She has also been active in career counseling in the Seattle area, particularly with respect to engineering and science careers for women. She is a consultant to UNESCO on the access of women to technical careers.

Economic optimization of energy conversion with storage (page 101)

Arthur Bruckner, II, received the B.I.E. degree in 1953 and the M.S.I.E. degree in 1958, both from the Georgia Institute of Technology. He served for three years with the U.S. Army Corps of Engineers during which time he was assigned to topographic mapping in Asia. After working as an industrial engineer for Sylvania Electric Products Inc., he entered engineering education as an assistant professor at West Virginia University. He later taught at Washington University and Northwestern University. In 1961 he joined The Boeing Company as an operations research analyst, where his research emphasis was on the design of information control systems for program management and on computer simulation studies of logistical problems. In 1965 he joined the Oklahoma State University faculty, becoming a candidate for the Ph.D. degree in engineering. His research work was in the area of power-plant economics. Since 1967 he has been with Louisiana State University, where he is an associate professor of industrial engineering.

Dr. Bruckner is a member of the American Society for Engineering Education, the American Institute of Industrial Engineers, and the Society of American Military Engineers. His current interests for sponsored research projects include the economic analysis of nuclear generation combined with energy storage and of wind-power generation plus storage simulation studies for underdeveloped areas.

W. J. Fabrycky received the B.S.I.E. degree from Wichita State University in 1957, the M.S.I.E. degree from the University of Arkansas in 1958, and the Ph.D. degree in engineering from Oklahoma State University in 1962. From 1954 to 1957 he served as a design engineer for the Cessna Aircraft Company. Summer faculty appointments include Chance-Vought Aircraft, the U.S. Air Force, the Ethyl Corporation, and Brown Engineering. He started his teaching work at the University of Arkansas in 1957. Later, at Oklahoma State University, he was an assistant and an associate professor. He is now a professor of industrial engineering at Virginia Polytechnic Institute. Dr. Fabrycky is the coauthor of three textbooks and a member of ASEE, NSPE, Operations Research Society of America, Institute of Management Sciences, American Institute of Industrial Engineers, Alpha Pi Mu, Sigma Tau, and Sigma Xi.

James E. Shamblin received the B.S. degree in mechanical engineering in 1954, the M.S. degree in mechanical engineering in 1962, and the Ph.D. degree in 1964, all from the University of Texas. At present he is an associate professor in the School of Industrial Engineering and Management at Oklahoma State University. His current research interests are in the fields of operations research and production control. From 1954 to 1955 he worked as a test engineer for Pratt & Whitney Aircraft. He then joined the Southwest Research Institute, working as a research engineer responsible for experimental tests and development.

In 1960 he began his teaching career as an assistant in the Department of Mechanical Engineering, University of Texas, where he subsequently became an instructor. In 1964 he joined the faculty of Oklahoma State University as assistant professor. He was appointed associate professor in 1966.

Game-theoretic applications (page 108)

James P. Dix received the bachelor of arts degree in mathematics from San Fernando Valley State College, Northridge, Calif., and also enrolled in the University of California, Los Angeles, for a year of mathematics courses. He served in the United States Army from February 1963 to August 1965. During this period he was assigned to Aradcom Headquarters, Colorado Springs, Colo., where, as a mathematician/programmer, he was responsible for design, analysis, and programming work on a wide variety of weapons systems. He was primarily involved with the Nike X system, utilizing the IBM 7094 computer. He joined the technical staff of the Control Data Corporation in 1965 as a programmer analyst and was later assigned to large systems. In his present position as supervisor of operations he has the responsibility for the hiring, training, and supervision of 28 data-center personnel.

Automation comes to the printing and publishing industry

Newspaper manufacture and distribution; communications

Revolutionary concepts and new techniques in printing, graphic arts reproduction, and lithography are literally catapulting this traditionally conservative industry from the "stone" age to the contemporary era of computer-driven equipment for the production and distribution of daily printed communication

Gordon D. Friedlander Staff Writer

The current information explosion not only has overtaxed our present-day library and archival resources—necessitating new concepts in information storage and retrieval systems—but also has severely strained our ability to produce the printed word and graphic art at a pace that will meet the ever-increasing public demand. Both the poor economics and low speeds of conventional typesetting and printing by hot metal type have militated against its continued use. We are witnessing a "revolutionary evolution" as new hardware and software are being developed at a breathtaking pace that will drastically affect the shape of the future flow of information and knowledge. At present, the nucleus of this effort is centered in the sophisticated general-purpose computers that drive electronic photocomposition machines and control press runs, subscription fulfillment, bundling, and distribution to spell out A-U-T-O-M-A-T-I-O-N for the industry. In the first installment of this two-part series, the automation of newspaper production, distribution, and communications is discussed. The concluding chapter, to be published next month, deals with the realm of magazines, periodicals, books, and specialized publications.

From the publication of the daily newspaper to the encyclopedia, industrial parts manual, department store catalog, and hard-cover book, the strong winds of change are sweeping the printing and publishing industry (hereinafter referred to as the "PPI") on a worldwide basis. Recently, the writer served as special consultant on the Committee on the Future of *The New York Times* to investigate the impact of technology upon the future of that great metropolitan daily.

Steadily rising labor costs, sharp drops in advertising revenues, inefficient throughput, and the resistance of the publishing trade unions—particularly in New York City—to automation in the manufacturing and distribution function of newspapers have resulted in the demise of four journals in that city during the past five years.

Actually, however, the experience of a number of leading newspapers, such as *The Washington Star*, *Wall Street Journal*, *Miami Herald*, *Los Angeles Times*, and others that have already automated their composing rooms, pressrooms, and mailrooms has indicated that few jobs have been eliminated; but the resultant economies and increase in productivity have been sufficient to tip the bookkeeping ledger from red to black ink.

The conventional printing processes

For those readers who are not familiar with standard newspaper printing processes, the following should serve as a brief introduction:

Typesetting. In letterpress printing, copy marked for size, type face, and line measure is sent to the composing room where it is set in type on one of the "hot metal"

linecasting machines—Linotype, Intertype, Monotype, etc. Foundry-cast type for handsetting is also used, but mostly for headlines and large headings.

The resulting composition, including heads, body copy (text), and captions—after proofreading and editing—is then assembled (dummying or makeup) into pages, along with necessary line engravings and photo halftones. In the printing of major newspapers, these pages are then locked up in heavy metal frames, called chases, which hold all of the page elements securely in place.

A thick sheet of papier-mâché (the stereotype matrix) is then placed over the page and squeezed down by high pressure so that, when removed and dried, it forms a type mold for the page. From this mat, a lead stereotype plate is cast in semicylindrical form, the curvature diameter of which adapts to the printing cylinder of the rotary press.

The rotary press. Large metropolitan dailies are printed on high-speed continuous-web rotary presses. Large, heavy rolls of newsprint are passed through the presses, which carry the stereotype plates on the surfaces of the printing cylinders. Each printing cylinder carries four to eight plates, thereby permitting four to eight newspaper pages to be printed on one side of the web at each revolution.

The offset press. In offset printing, thin flexible plates, made of zinc, aluminum, plastic, or plastic-coated paper, are used. The offset press contains four cylinders of uniform diameter. First is the printing cylinder, which carries the flexible printing plate clamped tightly in position. Bearing on the printing cylinder—one set on each side—are two sets of rollers: the water rollers, which moisten the face of the plate as the cylinder rotates; and the ink rollers, which supply the ink. Here, the "grease and water" principle applies: the water is retained on the nonprinting parts of the plate, but runs off the greasy printing image. The image, however, holds the ink (which is also greasy), but repels the water. The printing cylinder prints the plate image on the surface of the second, or blanket, cylinder, which is located below the first cylinder and tangent to it. The third, or impression, cylinder is located to one side of the blanket cylinder and holds the paper web firmly against the rubber blanket.

Printing takes place at the point at which the second and third cylinders are in contact. Thus the printing image is transferred, or "offset," onto the paper as the web is fed in between these cylinders. The fourth cylinder carries the printed web to the delivery end of the press.

Photocompositors. Photocomposition, or "cold type," machines electronically or optically generate alphanumeric characters in many type faces and fonts, which are then photographically imaged on film or paper.

The photocompositors are particularly compatible with the offset process, because the positive, right-reading images are immediately ready for photoengraving into offset plates.

Although letterpress rotary printing is still used for the majority of newspapers, the offset press is gaining ground, largely due to rapid advances in plate technology for the achievement of very long press runs.

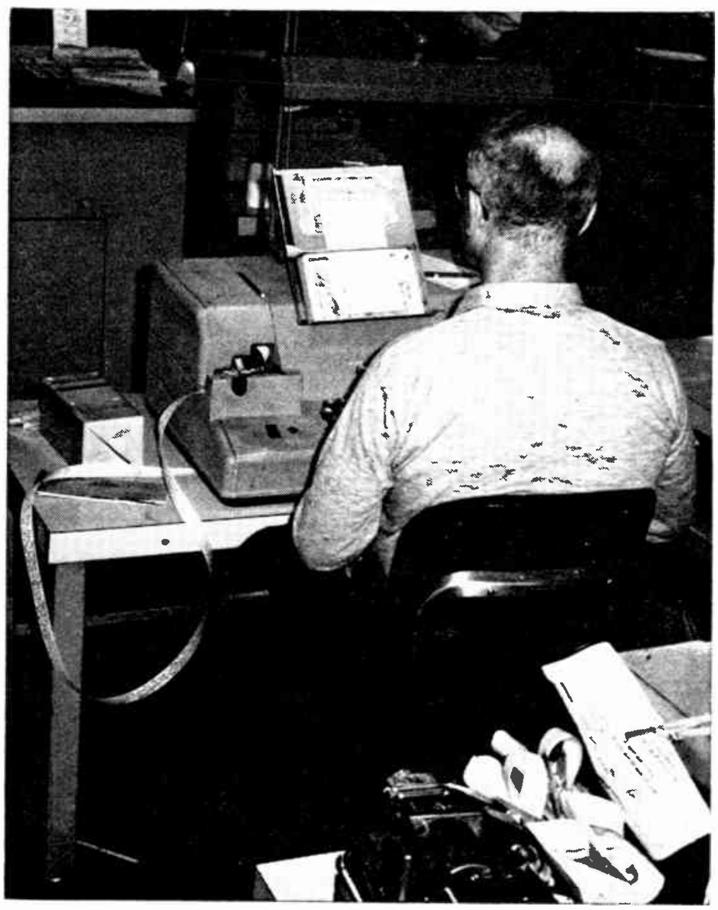
Some first steps toward automation in manufacturing

Essentially, the pathway from manual typesetting to the eventual complete automation of the composing room in a newspaper operation is traversed in a two-step procedure: first, the installation of a computer system with hyphenation-justification (H-J) capability that generates punched-paper tape to drive linecasters; then, the installation of high-speed computer-driven photocompositors (cold type) and cathode-ray-tube displays for graphic art reproductions and eventual light-pen editing as these systems become operational and available. In short, the first step may be considered as the short-term objective, and the second step as the ultimate objective to be completed from three to five years hence.

In this transition, it is important, of course, to select compatible hardware and software systems that can be updated as automated operations expand with the acquisition of additional modules and more sophisticated machines. Thus, the automation sequence in the composing room may be referred to as Phase I and Phase II.

FIGURE 1. The "heart" of *The Washington Star*'s automated capabilities—the central computer and data-processing center—containing two IBM 1800 computers, a Honeywell 200, and an IBM 1401.

FIGURE 2. One of the special keyboard machines simultaneously generating punched paper tape and hard-copy readout for the typesetting process.



Some examples of Phase I

In 1962, the RCA 301 solid-state computer was programmed, for example, for automatic hyphenation and type justification of editorial copy and classified ad "straight matter," with any of the commonly used punched-paper-tape-controlled linecasting machines (Linotype, Intertype, etc.). The system can also be used for paper-tape-controlled typesetting by photocomposition machines for display ads.

In addition to its typesetting capabilities, the 301 can be used off-line in three other principal areas:

1. General accounting, including payroll.
2. Subscription fulfillment, billing, and trend analysis.
3. Advertising scheduling.

Perry Publications, Inc. The 301 configuration is used by Perry, the publishers of the *Palm Beach Post-Times*, for typesetting in the following sequential steps of an input–output system:

1. Copy from reporters, rewrite men, and correspondents is edited at the copy desk; simultaneously, ad takers type classified ads as they are received.

2. The edited and corrected news copy and classified ads are retyped on a Friden Flexowriter (or similar machine) for the generation of the copy on punched paper tape.

3. Paper tape containing the clean copy is fed into the computer at the rate of 1000 characters per second. A seven-character message specifies any of ten type sizes and column measures that have been programmed for body text, cutlines, and classified and legal ads.

4. Paragraphs, up to 1000 characters long, can be read into the computer's magnetic core memory, with each word checked at the spaceband to see if the line is within justification range. Word spacing is attempted for justification of the line before hyphenation will be accepted.

5. If hyphenation is necessary, the word to be hyphenated is "looked up" in a 30 000-word dictionary that is stored on magnetic tape. If the word is not contained in this extensive storage, it is broken after the third, fifth, seventh, or ninth letter, and the line is filled with thin or "en" spacers.

6. When the paragraph is justified, it is punched on paper tape at the rate of 100 characters per second and simultaneously printed on the high-speed printer. Meanwhile, the next paragraph of unjustified paper tape is read into the computer.

7. The justified paper tape is fed into the linecaster (continuously, if for the same type size and line measure). Proofreading and corrections are handled conventionally. If a change in type or line measure is needed, a 30-second change is inserted into the computer's memory.

Los Angeles Times. The 301 is used in an alternative six-step sequence in which

1. The reporter writes his story on an electric typewriter, fitted with an attachment that simultaneously punches the copy on paper tape. He pencils corrections on the typewritten copy.

2. In the computer room, a typist cuts a paper "correction" tape containing the changes, deletions, insertions, and subheads put into the copy by the reporter and by subsequent editing at the copy desk.

3. The correction tape is fed, together with the original story tape, into the computer at a rate of 1000 characters per second. All the work necessary is done in the computer's memory in a matter of microseconds.

4. During the time interval between the continuous, 1000-character-per-second input rate and the 300-character-per-second output on paper tape, the computer automatically incorporates all editing changes and performs logical type justification.

5. The computer counts each character and keeps track of the number of units in each line. When it decides to hyphenate a word, it scans key-letter sequences to determine whether they follow the hyphenation rules for four basically programmed types. If not, certain letter sequences are analyzed to decide where the hyphenation should occur.

6. The justified paper tape is fed into a linecasting machine, which sets the type at the rate of 15 lines per minute. The type is then proofread and corrections are made; it is then ready for lockup.

By simplifying advertising detail and scheduling, the computer produces an ad manifest, listing the ad insertions by size for each section of the newspaper. The advantages of computer use for this function are that ad copy, type faces, etc., are compactly stored on magnetic tape; classified ads can be set, tabulated, inserted, and billed automatically; media records can be developed by classification and frequency for sales guidance; and greater speed is attained in makeup because large-size ads can be placed first.

In newspaper distribution, the computer can be used for the following major paperwork operations:

1. Maintenance of mailing lists.
2. Production of address and bundling labels.
3. Calculation of press run.
4. Analysis of subscriber and nonsubscriber lists and records (forecasting) for editorial, publisher, and circulation guidance.

FIGURE 3. General view of one of *The Star*'s four Linofilm photocompositors, showing the perforated tape drive and control panel. These machines were installed in 1959.

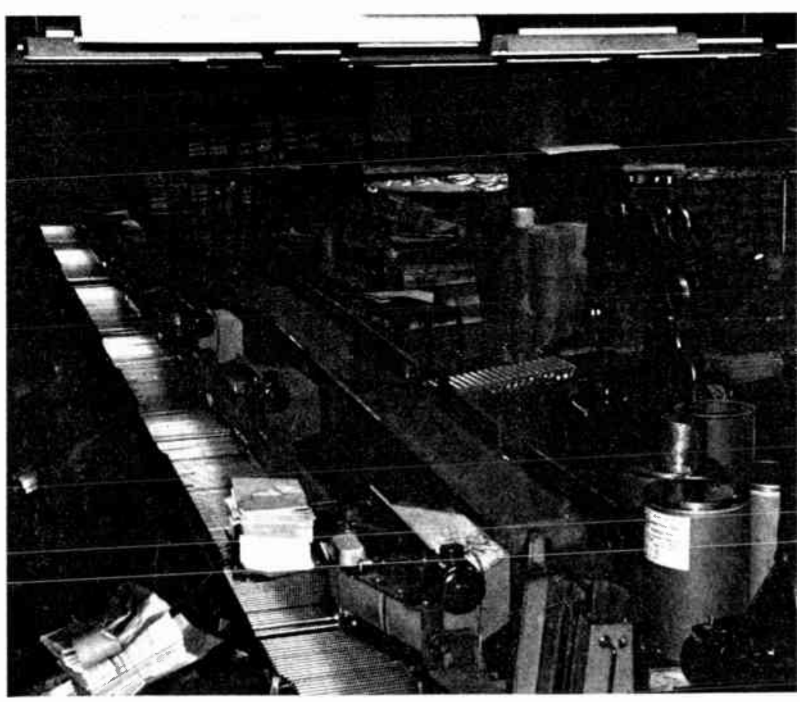
The Washington Star—modular design for the future

In 1959, *The Washington Evening Star* moved from its old building on Pennsylvania Avenue and 11th Street, N.W., to its new quarters—a building "designed for the future"—at 225 Virginia Avenue, S.E. This new structure was designed on a modular principle to permit future plant expansion with the anticipated growth of the newspaper. Thus, when future stages of plant construction are undertaken, they can be done with a minimum of dislocation to existing routines, production, and distribution flow. The mailroom, for instance, is designed for an ultimate 100 percent expansion.

The Star's plant is strategically situated alongside the Richmond, Fredericksburg & Potomac Railroad's right of way, and a private spur track into the building will simultaneously accommodate four boxcars for the unloading of newsprint. (Since the efficiency of an automation program involves a total concept for maximum speed and minimum lost motion in the throughput of a newspaper's news-gathering, manufacturing, and distribution processes, the interrelationship of the best available structural and mechanical design and planning—and transportation facilities—are prime considerations.)

The interior area of the unloading dock is paved to permit newsprint-carrying trucks to enter when the space is not being used by boxcars. A through truckway, adjacent to the mailroom, will permit the simultaneous loading of 18 newspaper delivery trucks. Storage space for a month's supply of newsprint is available within the plant. The 800-kg paper rolls are easily handled by a special

FIGURE 4. Bundled papers traveling along the mailroom conveyor system to any one of 18 preprogrammed down chutes to waiting distribution trucks.



WED 01/23/68

	NAME AND DROP POINT	ABC	YESTER DAY	TRAN	CHANGE	NEW DRAW	BUNDLE	ODDS	ROUTE	CARD SEQUENCE	NOTES
	OCEAN CITY TRUCK										
1	HOWARD JOHNSON RESTAURANT	36							318	10030	
1	HOWARD JOHNSON MOTEL	36							318	10040	
1	KENT ISLAND PHCY	36							565	10050	
1	RACK AT HOLLYS REST	36							638	10060	
1	RACK AT STUCKEYS	36							638	10070	
1	KENTMORR MARINA	36							638	10080	
1	CHESAPEAKE REST	36							638	10090	
1	BEACHCOMBER RESTAURANT	36							514	10100	
1	BILLS NEWS&VARIETY	36							662	10110	
1	TIDEWATER INN	36							662	10120	
1	REXALL DRUG	36							662	10130	
1	TRADERS PHARMACY	36							588	10140	
1	AYERS DINER	36							638	10150	
	AIR EDITION										
1		99							555	10170	
1	DAVE STEWARD-CAMBRIDGE	32							555	10180	
1	CARRIER SAMPLES	33							555	10190	
1		36							555	10200	
1		99							528	10210	
1	EMILE RAISIN-EASTON MD	32							528	10220	
1	CARRIER SAMPLES	33							528	10230	
1		36							528	10240	
1	BAY STATE NEWS-SALISBURY	99							662	10250	
1	BAY STATE NEWS-SALISBURY	32							662	10260	
1	CARRIER SAMPLES	33							662	10270	
1	BAY STATE NEWS-SALISBURY	36							662	10280	
1	LUTHER SHULTZ	36							658	10290	

FIGURE 5. An IBM printout giving the distribution and route schedule for daily newspaper delivery at the drop points indicated.

conveyor system and vacuum-fork-lift trucks to ensure minimum handling damage to the paper stock.

The "heart" of *The Star*'s automated capabilities (Fig. 1) is located in the new central computer and data-processing center. Here, all general-purpose and special-purpose computers and peripheral equipment (tape storage, card punchers and card readers, high-speed printers, etc.) are concentrated. At present, two third-generation IBM 1800 computers are replacing the older IBM 1401 and Honeywell 200. The new computers have comprehensive capabilities that include

1. The hyphenation and justification of all type.

2. The daily printout and updating of all distribution and circulation lists and truck routings.

3. Subscription billing, and general accounting.

Editorial and classified ad straight matter is typed on Flexowriters and Fairchild machines (Fig. 2) that simultaneously generate punched-paper-tape and hard-copy readout. This tape is then fed into the computer where it is transcribed onto magnetic tape. The magnetic tape is automatically processed and a new hyphenated and justified paper tape is generated to drive 14 Linotype Elektron hot-metal typesetters. Proofs are pulled for editing and corrections, and the corrected copy is reprocessed in the computer for the final punched tape to drive the typesetters.

Role of photocompositors. The same basic process is employed to drive the four Linofilm photocompositors (Fig. 3). These machines, installed in 1959, can generate 8 to 12 characters per second, depending upon type size, for display ads. *The Star*'s management is planning to update the now obsolescent Linofilm compositors either by the RCA Videocomp or Photon Zip machines (to be discussed in detail later in this article) that will be more compatible with the high-speed output of the new computer configurations. At present, 80 percent of the display ads are set by the photocomposition process.

At *The Star*, photographic film is preferred to photosensitive paper for use in the Linofilm photocompositors because the paper will not store satisfactorily for more than three weeks' time, and is subject to discoloration and fading.

Automation of the 'mailroom.' The distribution of newspapers has been greatly speeded by automation in the mailroom and by the adoption of new routing techniques. *The Washington Star*'s plant contains an excellent example of what can be done in a modern, automated mailroom arrangement.

The Star has an average daily circulation of about 340 000, and a Sunday circulation of approximately

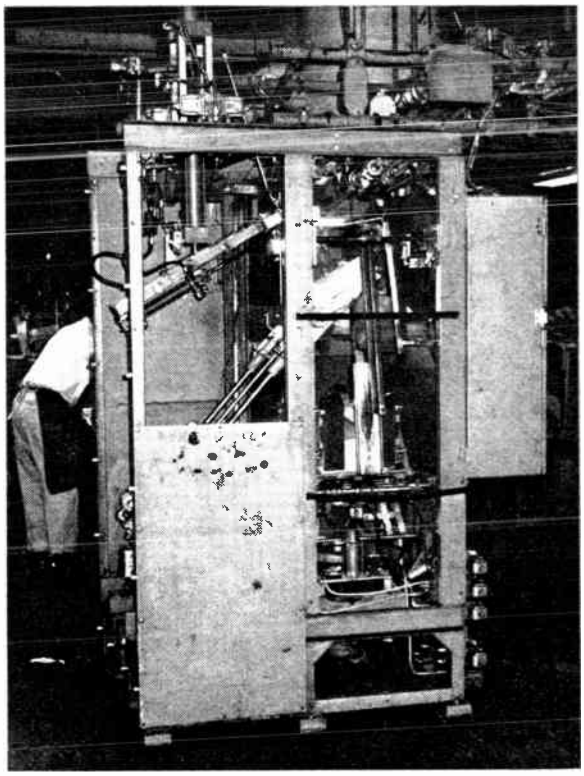


FIGURE 6. One of the automatic top-wrapper wire-tie bundling machines in *The Star*'s mailroom that "counts" the number of newspaper copies in each bundle according to the computer's program schedule.

FIGURE 7. The Sheridan console, which controls the proper number of newspaper bundles dropping down each chute to the trucks three floors below. This bundle count and chute-spotting operation is predetermined by the computer printout on the previous day.

390 000. From Monday through Friday there are five daily editions.

The mailroom is located on the third floor of the plant. The process flow is in a straight-line system, with 23 loading positions (Fig. 4), 18 of which are operative at any one time. Each day, an updated IBM printout (Fig. 5) gives the distribution and routing for the next day's papers. This encoded distribution printout also lists the drop points (city district, suburban town, county, and state), changes in draw and number of bundles, and the sequential number of drops in the assigned route.

Duplicate copies of the printout ticket are distributed to the console operator, whose controls guide the newspaper bundles to the proper trucks, to the truck drivers, and to the mailroom supervisor. This system completely eliminates the necessity for routing labels. Also, the computer estimates the number of pages for each edition, and the number of papers per bundle are controlled automatically by this computer estimate.

Forty-two Goss rotary press units run off the newspaper at 50 000 copies per hour. As the newspapers come off the presses, six Cutler-Hammer folders feed the papers into the top-wrapper wire-tie bundlers (Fig. 6), at the head end of six horizontal conveyors. When the papers (tied in bundles of 50, or more) come out of the automatic bundling machines, the operator on the Sheridan console (Fig. 7) takes over. A flashing light on his control console indicates that the proper trucks are in their correct loading positions as spotted and coded by the IBM routing printout. This light goes out when the truck driver in each bay pulls a collapsible roller conveyor from his truck and connects it to the loading chute.

The console operator then pushes a button and the newspapers move out of the bundling machine and travel along the conveyor to the correct chute. The operator sets a dial for the proper number of bundles to be sent down the first chute. A deflector pushes the bundles to the proper chute, and the flow is automatically cut off when the correct number of bundles drop down that chute and into a waiting truck. When the automatic counter reaches the same number as set on the console dial, the deflector automatically retracts and the bundles move to the next down chute, at which point this loading procedure is repeated.

When a light flashes in a loading bay, the driver of the particular truck is warned that the last bundle for his delivery is coming down the chute. Thus, by this automatic counting and control, the truck driver does not have to record the number of newspaper bundles; his copy of the IBM printout contains the number aboard, their destination, and the best routing for his truck. As soon as one loaded truck leaves its bay, another truck—waiting in queue—pulls into the vacant position for loading.

The supervisor of this operation estimates that the automatic equipment has reduced the number of mailroom personnel by at least 22 men. The total number of *Star* employees, however, is about 1700—a figure not much below preautomation days. The paper's management makes every effort to reassign personnel displaced by machine technology.

The overall impression of *The Star*'s operation is that of a compact, efficient, and smooth-flowing system in which there is no apparent lost motion or duplication in operation. All processes seem to be set up in straight-line flow patterns from the raw material stage to the

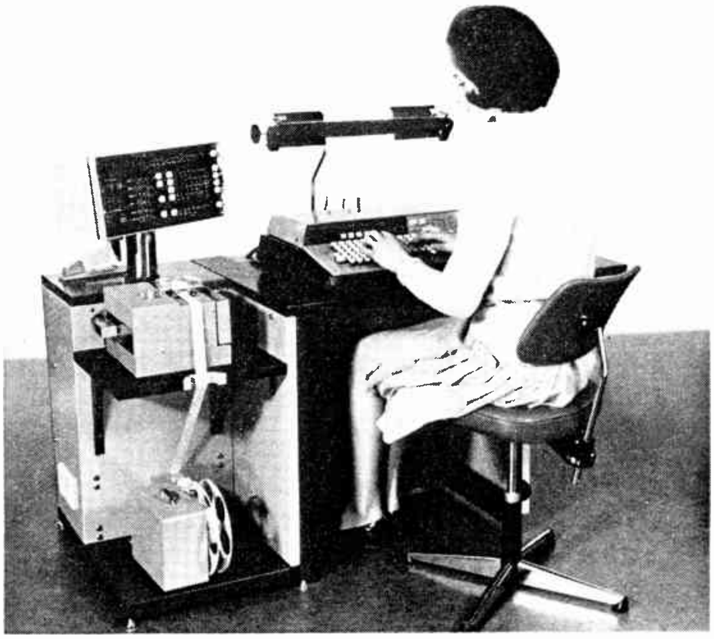


FIGURE 8. The 505 keyboard unit containing an electronic perforator that features a rapid interchange of keybanks, from the standard 64-key layout to the extended 78-key set and the 90-key Linotype layout.

manufactured product. The composing room, pressroom, and mailroom are clean and uncluttered.

Development of photocomposition machines

In the beginning, with first-generation photocompositors, character generation was a relatively slow, letter-by-letter process. In the second-generation machines such as the Mergenthaler Linofilm and the Photon 713, first produced about 1957, the character generation rate was stepped up to about ten per second, depending upon the type size and font required. Then, in a transition between the second and third generation, a custom-built Photon machine called GRACE--an acronym for *GR*aphic *A*rts *C*omposing *E*quipment--was installed three years ago at the National Library of Medicine in Washington, D.C. Computer-driven, and operating at a character generation speed of about 300 per second, GRACE can use 226 different characters in preparing a 23-cm-wide positive photographic film. This photocompositor, which is marginally compatible with computer speeds, will be discussed more fully in the context of the concluding installment of this article.

The RCA Videocomp

The Videocomp is an all-electronic third-generation photocompositor, which operates at computer-compatible speeds.

Essentially, the machine consists of four basic components: a computer core memory, a digital control, an analog control, and a high-resolution cathode-ray tube that is capable of generating up to 650 letter characters per second across its face and up to 900 lines per minute (depending upon the line measure). At this rate, the text for a complete newspaper page could be composed in about two minutes' time. The Videocomp can "set" type, by means of electronic dots and scan lines, in two size ranges: 4 to 12 points, and 8 to 24 points.

In ideal operation, a computer (preferably an RCA Spectra 70, which is well adapted for use with the Videocomp) would automatically hyphenate and justify keyboarded input copy and provide all the necessary instructions on font preparation by digital encoding to the Videocomp tape reader. The letter characters are digitized into stroke segments and placed on magnetic tape. This is the input information that contains the fonts to be used. The tape is read electronically by the Videocomp, which calls from its memory the proper characters in the desired type font and point size. The letters are flashed onto the CRT and are sequentially exposed, by means of a precision lens, onto sensitized paper or film of reproduction quality. The paper is developed in a one-pass operation, and the output is now ready for paste-up in the dummying or makeup of a newspaper page. The film is developed off-line in cassettes.

The technique is so flexible that any letters—from Chinese calligraphy to the Cyrillic alphabet and Old Testament Hebrew—can be stored either in the internal memory or on external tapes for instant access and use. Scientific and mathematical symbols and engineering notation present no problem in setting. It can also set new fonts within a line, handle subscripts, superscripts, and foreign language diacritical marks, vary line lengths to handle "runarounds" for accommodating illustrations, and compose text in column widths up to 31 picas.

The type faces can be electronically manipulated into italics, condensed, or expanded type to meet the esthetic or visual requirements of display ads and special text messages. Up to four fonts can be stored in the machine's internal memory, and a virtually unlimited number of fonts can be stored externally, either on paper or magnetic tape.

Each Videocomp character is composed of more than 2000 electronic "strokes," thereby giving a resolution quality comparable to foundry type. Practically all the operations of this electronic photocompositor are automatic. The only moving part is the film or paper advance. All circuitry is solid state. According to RCA, the machine is easy to operate, and one man--trained in a matter of days—can handle the device.

The Linotrons

The Mergenthaler Linotype Company, in collaboration with CBS Laboratories (the subcontractor for electronic R&D and the manufacture of hardware components for Mergenthaler), has produced the first operational configuration of the Linotron 1010 for the U.S. Government Printing Office, and a smaller version, the 505, which will be commercially available this spring.

Linotron 505. The basic configuration for the 505 is modular and consists of three components: the control unit containing integrated logic circuitry for system control, plus a central character width core storage for four grids; a reproducer unit, which houses the character grids, the character generators, an index CRT and printout tube, and film magazine; and an electronic keyboard, which can address the system by means of eight-level perforated tape. An optional tape-merging unit can also be provided for correction tape.

Complete newspaper requirements can be met by the 505; it has the full capability to handle the most complex work, such as newspaper display advertising. Leading and letter spacing are controlled from the tape coding. A resolution mode of 1300 scan lines per inch gives a very-high-quality typographic resolution at a rate of 90 newspaper lines per minute. In the standard resolution mode of 650 scan lines per inch, however, quality typography is produced at the rate of 120 newspaper lines per minute.

Although the normal input to the control unit is eight-level tape—with six levels used for coding, one for the shift function, and one for parity—the flexible control unit is not locked into any particular keyboard or computer, and may be programmed to accept six-level TTS (Teletype system) tape or, with the proper interface, magnetic tape.

Machine operation. The 505 keyboard unit (Fig. 8) contains an electronic perforator that features a rapid interchange of keybands—from the standard 64-key layout to the 90-key Linotype layout.

The 505 control unit receives the input at the punched tape input station by means of a photoelectric reader. For magnetic tape input, a suitable magnetic tape reader can be added.

A central character-width store within the control unit is divided into two parts: one holds the information for the four grids in the reproducer unit; the other covers the requirements of any 12 additional grids. A multiplexer allows for full use of up to 16 keyboards that may be connected to the control unit (by cable), on a time-sharing basis, without the necessity of holding character-width information at each keyboard station.

Type images are formed by a combination of grids and CRTs in the reproducer unit. These images are built up by successive lines on the printout CRT, and these are projected onto film by a traversing lens—precisely controlled by means of a fine-line grating—that sweeps across the required line length. Each grid (Fig. 9) contains a total of 256 characters, but since 16 locations are used for control functions, the net total of available characters is reduced to 240. The grid change mechanism can be loaded with any combination of four internal grids to provide an on-line typographic capacity of 960 characters, with a full range of 15 point sizes ranging from 4 through 30 points, and with 3, 36, and 42 points as options. Access to a maximum of 12 external grids (for a total of 2880 characters) can be achieved, and the selection of these grids is automatically controlled by tape codes.

A raster, or television scan-line screen measuring approximately 3.8 by 3.8 cm, is formed in any one of the 16 control positions on the first CRT (index tube) by means of "flying-spot" scanning. Sixteen corresponding lenses focus the light from the scanning lines onto one character in each of the 16 control positions on the grid. Behind the grid in use is a set of 16 photomultipliers, only one of which will be in operation at a given time; thus one character is selected for reproduction.

The reproducing system consists of a second CRT (printout tube), plus a moving lens and a film magazine. The light spot generates straight lines on the face of the printout tube in synchronization with the successive raster lines. Light from the tube is collimated and passed to the lens by means of a mirror. The lens forms a reducing image (20:1) of the CRT face on the film while sweeping across the required length of the line.

The previously mentioned fine-line grating, attached

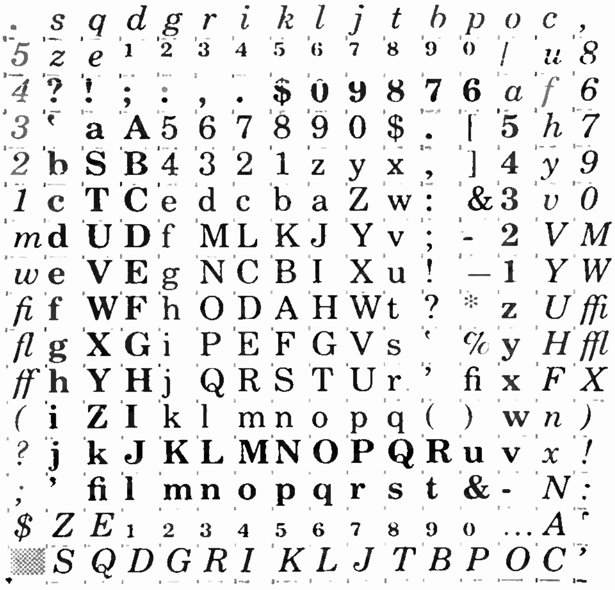


FIGURE 9. Typical Linotron character grid, containing 256 alphanumeric characters and standard symbols in roman, italics, and bold face.

to the lens carriage, interrupts a beam of light between a static light source and a photocell, thereby providing 650 to 1300 pulses per inch (2.5 cm) of carriage movement. These pulses form the "master clock" for the entire system, since the lines of the flying-spot scanner and the printout tube are synchronized with them. Also, the pulses are counted to give accurate positioning control of the printout.

The film, or photosensitive paper, advances to the position of the next line during the return movement of the lens carriage. This advance, corresponding to the leading of hot-metal casting, is controlled by the input signals in increments of one-half point. The cassette will hold a length of about 33 meters of film or paper.

Linotron 1010. The "big daddy" of the Linotron[1] family is the model 1010, the electronic composing system that has been installed at the U.S. Government Printing Office in Washington, D.C. The configuration consists of

1. A general-purpose computer that can be programmed to accommodate a broad range of composition instructions.
2. An electronic phototypesetting machine.
3. The necessary auxiliary equipment, such as film processors, tape converters, and input keyboards.

The electronic phototypesetting machine (photocompositor) elements include the control unit, display unit, and character generator. It is capable of generating, displaying, and positioning any symbol—in random order—and as specified by a control and information magnetic tape that has been generated for it by the computer.

The machine can set a full-page newspaper makeup at one-quarter size (about 20 by 26 cm), in which the full-face display area of the CRT is utilized. Thus the 1010 has the ability to set an entire newspaper page instead of the column-by-column capability that limits other automated compositors. Enlargement to full-size newspaper page is accomplished by off-line optical means. According to Mergenthaler, the machine can set nine full columns of a newspaper classified ad page in 50 seconds' time.

Equipment speed. The speed range will vary with letter character size. For example, at 6 points (standard newspaper classified advertising size), the speed will be 1000 characters per second; at 8-point size (standard newspaper text), 800 characters per second may be generated; and at 10-point size (normal book text), the speed of generation is 620 characters per second.

As a further indication of the machine's composition speed, a 4000-character page, set in 8-point type, will be produced in less than six seconds—or at the rate of approximately 600 pages per hour, including the required one-half second to advance the film at the completion of each page.

Character grids. The Linotron type fonts are etched onto glass plate grids, each of which contains an array of 256 characters. Four of these grids are mounted on a circular disk (Fig. 10) that may be rotated to swing any of the four grid plates into position, in one-half second, for scanning as required. Thus more than 1000 characters are available on each disk. The disks are stored in a circular magazine similar to that of the Kodak Carousel slide projector. Characters are available in eight different point sizes, ranging from 5 to 18 points.

A standard character grid (or job set), which is comparable to conventional job sets now used in the PPI, has a complete font of type that includes alphabetic, numeric, punctuation, and related graphic arts characters. The grids are also capable of producing characters in roman and italic styles, small caps, and boldface—plus the capability of producing subscripts and superscripts, arithmetic fractions, and foreign language accent marks.

Principles of operation. In the Videocomp, the characters are stored on magnetic tape in a computer memory, recalled when needed, and then electronically generated. The Linotron 1010 principle, as we have seen, is an optical-scanning process from master photographic plate grids. Further, the Linotron readout is not on continuously processed photosensitive paper; it is on direct film negatives that are housed in cassettes and then developed off line.

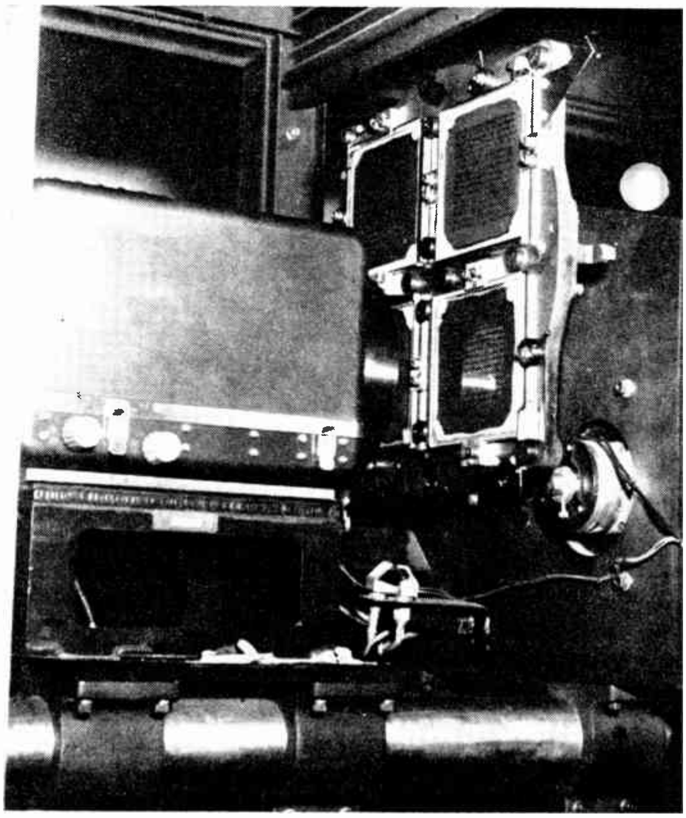
The 1010 equipment is housed in four separate cable-connected cabinets: the first contains a magnetic-tape reader and a conventional magnetic-tape drive that is used by the computer; the second encloses a tape-reader control and the control electronic circuitry; the third module contains the character generator; and the fourth contains the page-display CRT and film carrier.

The character grid is optically projected (Figs. 11 and 12) to the face of the character-generator tube, and the photocathode face of this tube converts the light image of the characters into electrons. Vertical and horizontal selection bars receive a signal from the control circuits and a single character emerges as a video signal. Servo controls also receive information from the character generator to assist in the precise positioning of the character on the CRT.

The complete process by which a desired character may be taken from the magnetic tape, through the photocompositor, to the final film output is accomplished by the following sequential steps:

1. The operator pushes the start button and up to 2048 data and control characters (from two disks) are read into the control buffer from the magnetic-tape reader.

2. The control circuitry reads the machine function codes, which indicate the point size or grid number, and the first character is then ready for selection.



3. The position of the first data character, in relation to its place on the film page, has already been coded on magnetic tape by the computer. The character code is decoded, and a signal is sent to the character generator to open the gate for that character and let it emerge as a CRT display.

4. The television signal for the character is sent to the display CRT and positioning information is transmitted to the servo-control CRT. As previously indicated, the servo tube and its circuits position the character precisely and then transmit control information to the display CRT to display the character and, simultaneously, to project the image of this character through lenses to the photographic material.

5. As soon as a character is scanned, the next character in the word is released by the same process until an entire page is composed and a film advance signal is received to start a new page.

Master composition program. The 1010 will not perform any operations that have not been written into the master composition program and recorded in the memory of the general-purpose computer. This program encompasses a complex series of routines that provide the computer with instructions to perform every operation that the compositor must do—from setting galleys of type to page makeup.

In preparing various styles of composition, the master composition program must be furnished with specifications for each individual job. These job instructions, which are comparable in all respects to those that must be given to a human compositor, include type face, type

FIGURE 10. (Left) Four character grids, mounted on a circular disk turret in the 1010 configuration, may be rotated to swing any of the four plates into scanning position in 0.5 second's time.

FIGURE 11. (Right) Isometric diagram illustrating the manner in which high voltage applied in one horizontal and one vertical switching bar in the Linotron 1010's character generator unit selects a beam which then carries the video signal for a single character. Thus the left and right beams (shown as colored arrows) are blocked, and the center beam is collimated and transmitted.

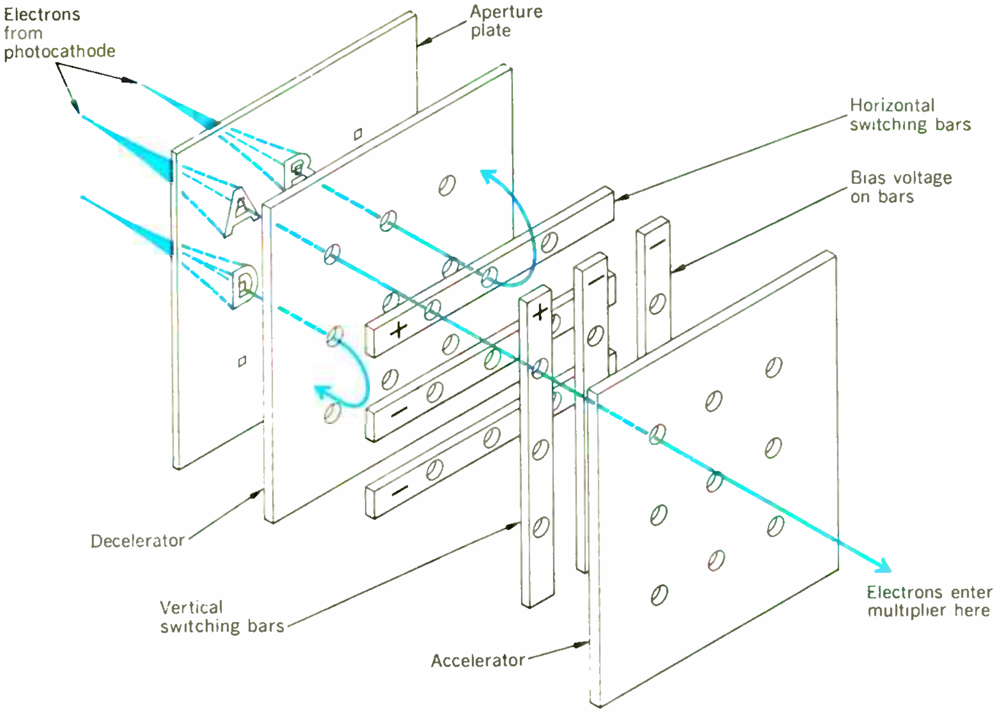
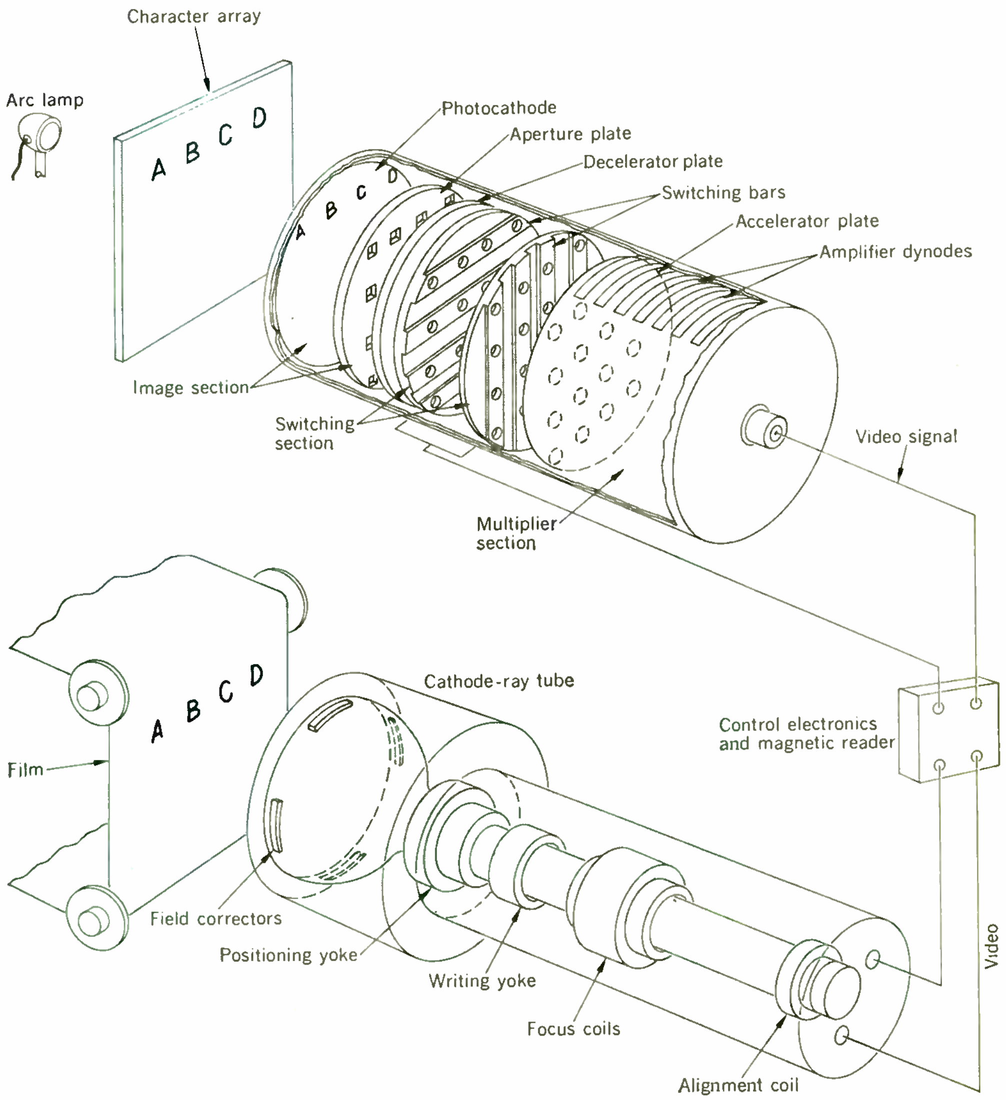


FIGURE 12. Input/output diagram showing how the character generator's switching section selects one of the 256 video signals (representing the grid characters that have been projected on the tube's photocathode). The video signal is then sent to the CRT where it "sweeps out" the character in its proper position on the page to be printed. The characters expose film which is developed off-line and used to make printing plates.



size, leading, line measure, page width, page depth, running head location, and folio location. This is essentially the page format information.

As a page format is developed for each new application, that format will be identified by a code number and placed on a parameter tape file from which it may be recalled for future requirements of a similar nature. At present, the master composition program for the Government Printing Office has been written for an IBM 360 computer. This program will later be adapted to meet the requirements for other computer systems as the need becomes apparent.

Photon ZIP 901

The ZIP 901 consists essentially of three elements: the *electronic control console*, which processes information received from the standard six-bit *magnetic tape handler* (furnished by others), and then passes these data directly to the *photographic output unit*.

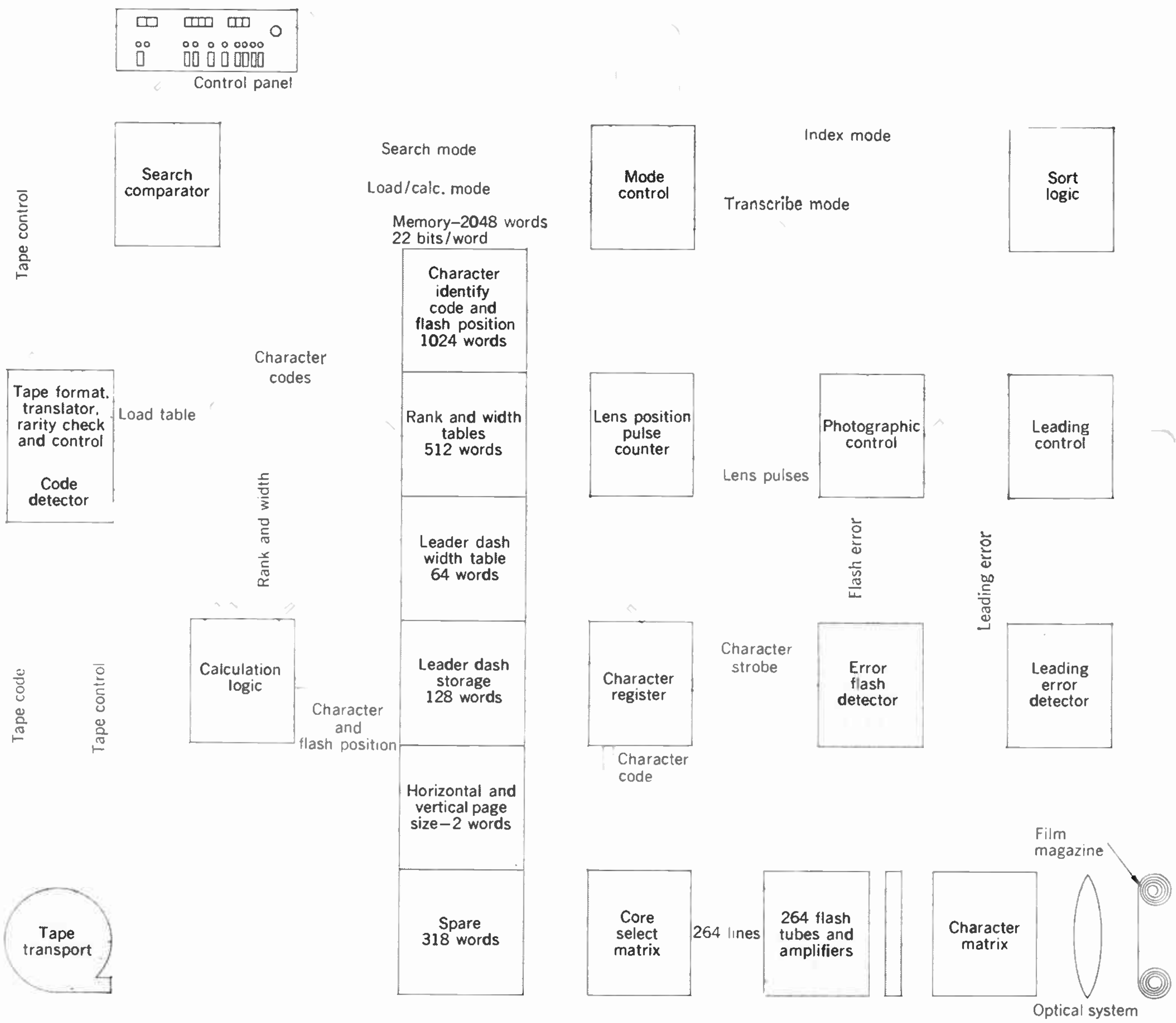
The control console provides all necessary manual and automatic control functions and indicating devices to achieve the desired operation. Thus the information input to the console contains a sequence of characters, presented in the order in which they appear in the printed line, together with control information that gives the location of the line in reference to the text to be printed.

The control console incorporates for major subsystems:

1. Control panel—required by the operator for system operation.
2. Maintenance panel—comprising a set of controls and indicators to ensure proper function and to simulate test conditions.
3. Memory unit.
4. Solid-state logic.

The 901's photographic output unit contains the mechanisms and solid-state circuitry necessary for the acceptance of signals from the control console (see Fig. 13 block diagram).

FIGURE 13. Block diagram of the Photon ZIP 901 photocomposition system.



Basic operation. Eleven rows of 24 negative (transparent) character images (Fig. 14) are contained in a stationary glass matrix. Each of the 264 alphanumeric characters of the matrix is provided with individual xenon-flash illumination. Character images are reproduced on film or paper by a lens that traverses continuously back and forth between two fixed positions in such a manner that the entire film width is scanned. A memory storage system and controlling circuits produce timing for the illumination of individual characters.

Output speed is a function of line length, not the number of characters in a line. It varies according to type style and size; for example, using 6-point type (30 characters per inch) at a speed of 2.37 nine-inch-long lines per second, the composition rate is 640 characters per second. Speeds of more than 750 characters per second usually can be achieved by the use of more condensed type faces.

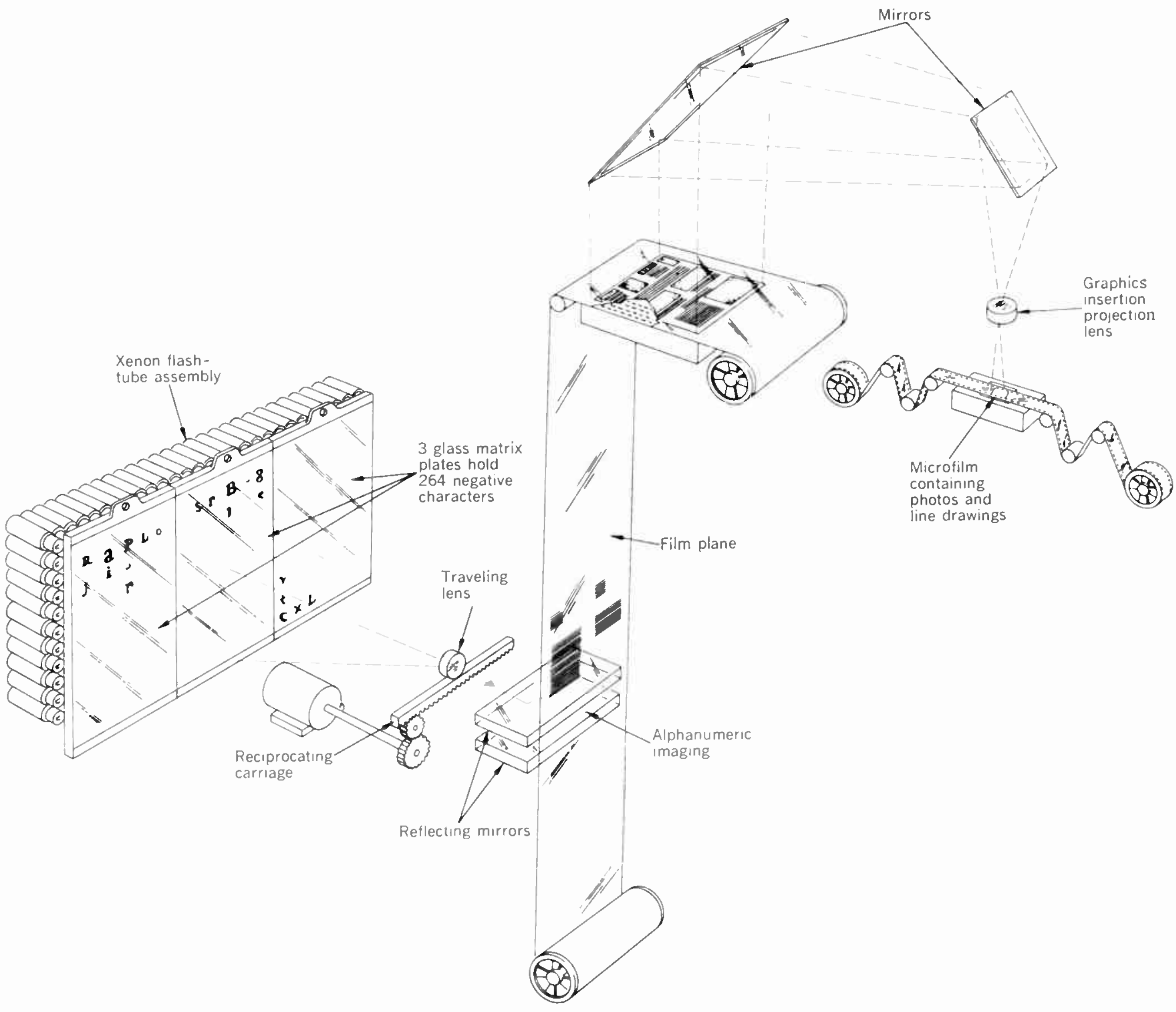
Optical system. Each xenon-flash lamp is operated by the discharge of a capacitor, which is triggered by a pulse generated at the appropriate instant. The lens is positioned between the matrix and the film so that, at any instant, it can make an image on the film of all characters in the matrix. The lens is mounted on a carriage (as shown in Fig. 14), which reciprocates between two extreme positions. A full line is created in each traverse by firing the appropriate combination of lamps.

The displacement of the film for line spacing occurs during the reversal of the lens movement in less than 90 ms. A grid attached to the lens carriage, an optical system, and a photocell generate a timing pulse for each unit displacement of the carriage. The film cassettes are designed to accommodate 40-meter lengths of film, up to 30 cm wide.

News gathering and communications

In the modern newspaper plant of today and tomorrow, the latest electronic equipment for the news-gathering and

FIGURE 14. The simultaneous imaging and page composition of copy, photos, and line drawings can be achieved by the 901 computer-driven electronic photocompositor.



communications function is as important to the primary input phase of newspaper operation as the manufacture and distribution processes are to the output end of the continuous cycle of journalism. Here, the writer will present a brief overview of a novel system for remote data acquisition and its transmittal to a data terminal—in this case, the home office of a newspaper— and discuss the basic elements of facsimile transmission.

The Data-Verter system. The Data-Verter developers (Digitronics Corp.) are primarily concerned with data acquisition and the transmission of these data from outlying stations, in computer-compatible form, to a central data-processing point. The Data-Verter consists of five basic components: a special typewriter, a digital recorder, an acoustic transmitter, a data processor, and a Flexowriter (or equivalent) printout. The first three components are located in the remote station, and the last two are in the central office.

In operation, a reporter in the remote office would type his story on the input machine, which would simultaneously produce a hard-copy printout for the correspondent's files and record his story, in digital form, on magnetic tape contained in a cartridge within the digital recorder unit. When the complete story is recorded, the reporter simply removes the tape cartridge and places it in the acoustic transmitter unit. He then dials a coded number to his home office to inform that he is ready to transmit a story.

By placing the telephone handset in a special cradle of the acoustic transmitter, the correspondent can then send the encoded data over standard telephone lines, WATS, or private lines at the rate of 360 words per minute. At the data center, a punched tape is produced from the coded message that is received. By taking this tape and feeding it into a tape-driven typewriter, the reporter's story, exactly as he composed it, will be recreated as the printout hard copy on the home office machine.

Facsimile copy. Facsimile transmission is being used to some extent at the present time for sending full-size proofs of newspaper pages over long distances. This is done at high speed and with precise definition.

In operation, conventional proofs are placed on the drum of a transmitting console. When the unit is switched on, the drum begins to rotate and the pages are scanned by a precise optical system. The light from this system is converted into an electric signal, which may be transmitted either by microwave or over standard telephone circuits. When received by the recording equipment, located at a remote point, the electric signal is reconverted to light and scanned on photosensitive paper sheets on the recording drum. The resolution of the reproductions thus made are usually so good that the naked eye cannot distinguish between the original and the copy.

Photo engravings are made directly from the duplicate pages, and matrices are then made directly from the engravings. Although letterpress reproduction is generally used, photo-offset plates can be made directly on the facsimile recorder drum. In newspaper work, screened halftones transmitted on a proof page will produce screened engravings without loss of quality.

Facsimile systems generally provide a scanning capacity range from 300 to 1500 lines per inch for the accurate reproduction of typographic detail. The transmitters and recorders use synchronous motors, without gears, to drive the transmitting and recording drums. Precise transmission frequency standards (with a maximum bandwidth of 4 MHz) and control signals lock the scanning and drum-driving systems in synchronism at both the transmitting and recording units for accurate tracking. These control measures reduce or eliminate such distortions in facsimile transmission as "jitter" or "gear pattern," and other irregularities.

The transmittal (and recording) time for a full newspaper page, at a scanning resolution of 800 lines per inch and a drum speed of 3000 r/min, is about four minutes.

The manufacturers of this equipment emphasize that it is "engineered for the future," since it is adaptable to communications systems that use earth-orbiting satellites. Newspaper pages have already been transmitted experimentally via Telstar.

'Intercontinental' typesetting

Almost five years ago, RCA and United Press International (UPI) achieved an historical first in printing, by demonstrating that an electronic computer, located in one continent, could transmit instructions to automatic linecasting machines in another continent. This demonstration was carried out from Chicago in connection with the 35th Annual Production Management Conference of the American Newspaper Publishers Association. News copy originating at the ANPA convention was transmitted by conventional Teletype circuits to Camden, N.J., where it was fed into an RCA 301 computer. The computer automatically hyphenated and justified the copy, then sent instructions via RCA Communication's submarine cable to keyboardless Intertype linecasters at the *Manchester Guardian*, *Glasgow Herald*, and *Edinburgh Scotsman*.

Shortly thereafter, the Relay I satellite was used successfully to transmit news copy from UPI's Rio de Janiero bureau to Camden, where the RCA 301 processed the copy and forwarded it to Chicago. There, it was fed into UPI's national teletypesetter circuits, which serve hundreds of American newspapers. In a parallel experiment, Associated Press (AP) sent news copy by cable from London to New York and Chicago via the RCA 301 at Camden. And, on June 14, 1963, RCA and NASA carried out an additional experiment in which computerized news copy was sent to British newspapers by means of Relay I.

The 'WSJ' and the 'satellite plant' concept

Several years ago, Dow Jones Inc., publishers of *The Wall Street Journal*, anticipating protracted difficulties with the labor unions if it centralized the bulk of its publishing operations in New York City, methodically proceeded to decentralize. Today, it has eight publishing plants throughout the United States, and it is the second largest daily newspaper in the country, with more than one million circulation.

The most recent step in updating its printing and publishing techniques has been the establishment of a photocomposition and offset printing plant at Highland, Ill., where news stories and ads are composed on film that is used in making offset plates for printing the *WSJ* for a portion of its Midwestern Edition readers.

On February 27, 1967, the *WSJ* readers in northern California and the Pacific Northwest received their first papers from Dow Jones' Palo Alto plant, the newest of a string of printing centers. This facility is one of the most modern in the United States. It contains electronic typesetting equipment, a five-unit press capable of producing 70 000 copies of the *WSJ* an hour, and automated mailroom machinery. Here, facsimiles of each page of the Pacific Coast Edition are transmitted 640 kilometers south, via AT&T microwave (utilizing a full television channel) and by a parallel telephone Class "A" line of 3-MHz bandwidth (to ensure a "clean line" in which there is minimum possibility of facsimile distortion by echo signals during transmission), to another *Journal* plant in Riverside, Calif. At Riverside, the facsimile images are etched onto metal plates from which the *WSJ* is printed for readers in southern California and neighboring areas. The Pacific Coast Edition has a circulation of 171 000, of which about 104 000 copies are printed at Palo Alto and the remainder at Riverside.

At the present time, Dow Jones uses two AT&T news circuits, but preparations are under way for the installation of a central computer telecommunications system that will link, via telephone lines or microwave, all eight *Journal* plants and New York City. Two IBM 360 computers have been installed, for this purpose, at the South Brunswick, N.J., plant, and full operation of the extensive network is scheduled for this year.

Other computer and photocomposition installations at *WSJ* plants include two IBM 1130 machines at Chicopee Falls, Mass., to take over the task of hyphenation and justification of news copy for typesetting as part of a nationwide electrotypesetter system. A Videocomp has been installed in the South Brunswick plant, and an RCA 301 data processor at Chicopee Falls is used for record-keeping purposes and to handle 1.6 million subscription accounts. Names and addresses are stored on magnetic tape, and the machine has random-access capability for continuous updating of subscription status.

The newspaper of the future

The short-term outlook. Insofar as newspapers are concerned, the next five to ten years will witness increasing automation in the composing rooms and mailrooms at existing newspaper plant. Although labor resistance, in many instances, will be the determining factor of the rate at which hyphenation-justification computers, automatic linecasters, centralized multipurpose computers, and mailroom consoles can be phased in, this equipment will eventually be accepted and used.

The format of most major newspapers will probably change little in the next five years or so, according to the American Newspaper Publishers Association. The current research philosophy of the ANPA, however, is in the direction of satellite plants and offset printing. This trend is based on the reality that the full capability of computer-driven typesetting equipment cannot be achieved with hot metal and the conventional stereotype rotary printing process. To attain the full potential of computer speeds, the machine must be coupled to a photocomposition system that is capable of a character-generation rate of 700 to 1000 per second; the film or photosensitive paper output of which is camera-ready for final etching onto the offset cylinders.

Nevertheless, the ANPA is also conducting research in letterpress printing techniques, including the development of plastic stereotype plates to replace lead. If a genuine breakthrough in letterpress methods is achieved, the association will redirect its efforts in that direction.

At present, the ANPA is involved as part of a joint committee with the Bell System in a study that includes:

1. Ways in which to apply modern communications technology in press communications to improve the speed and efficiency of such communications.
2. Methods to help improve the day-to-day newspaper administration, operations, and marketing activities.
3. Identification of what communications requirements are at present and in the future, and what new developments should be explored.
4. Application of information storage and retrieval systems to the future needs of newspapers.
5. Evaluation of the performance of press communications services.

The association is also very much concerned with finding processes to enable newspapers to use the electronic computer and computer-driven typesetting and photocomposition equipment without causing the obsolescence of existing equipment and the attendant financial dislocations to the newspapers involved.

One of the overriding reasons that ANPA is directing its research energies toward offset printing and satellite plants—in addition to the adaptability of offset printing to computer speeds—is the simplification of nationwide distribution made possible through the establishment of such plants in strategic locations. Also, since the press run by the offset process is generally not as high as with the letterpress process, a larger number of satellite plants would probably be required for high-volume production. The use of facsimile transmission would be an essential adjunct in such an operation.

Long-term perspective. Over the long term (10–20 years hence), newspapers will probably change considerably in their format and readership patterns. The newspaper plant of the future, in its editorial and graphic arts functions,[2] will probably have complete electronic write–edit–compose facilities (see Fig. 15). For example, the writer or reporter might work on an incremental magnetic-tape typewriter, such as IBM's new Selectric cold-type system. This would record copy in digitized coded form as it is written, including changes and corrections made by the writer as he proceeds. Raw copy tapes would then be transmitted to the editor (either by telephone lines or microwave, depending upon distance).

The taped copy would be scanned by the computer to eliminate discarded material after the author had made his corrections. Following this step, the clean copy would be displayed on a CRT, and then the editor would make changes directly on the face of the tube by means of an electronic light pen and keyboard built into the display unit. This editing would be instantly recorded in the coded manuscript. The editor could also instruct the machine to prepare automatically any indexes, tables, or abstracts that could be handled by preinserted machine programs.

Art directors and illustrators will probably work on the face of CRT displays rather than on paper. Although final artwork may still be prepared by the use of conventional materials, and inserted into the system by an electronic scanner, the CRT display unit would be excellent for sketches and preliminary studies. The artwork could be displayed simultaneously in several of the newspaper's editorial offices, discussed by telephone, and revised on the spot by any of the participating editors.

Further, the art director could insert graphic specifica-

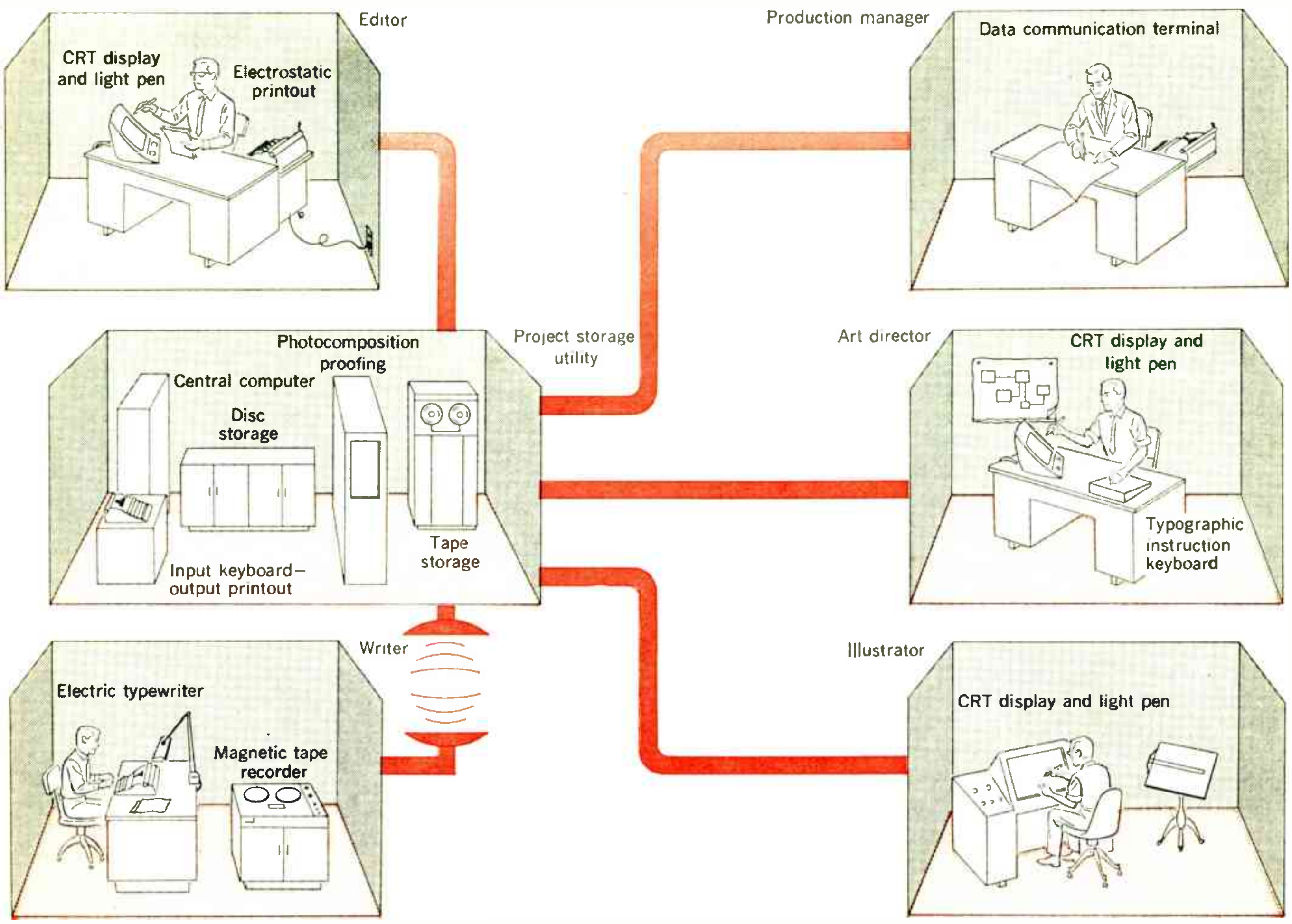


FIGURE 15. Diagram of the complete electronic write-edit-compose capabilities of the future.

tions into the system from his console for accurate page makeup and could receive proofs for verification within seconds. He could then either view these proofs on his display screen or obtain a hard-copy printout from a character-generating proving unit.

The production manager could employ a teletypewriter data communication unit terminal to insert appropriate data and instructions in order to receive various management reports such as cost projections based on current character counts, graphics, design specifications, or job time schedules on file within the computer system. Also, he could call for the automatic transmission of purchase and production orders.

In this long-term area, audio response, or voice recognition, will be a reality. Today, the New York Stock Exchange utilizes a device with a vocabulary of about 32 words to give current stock prices from a continually updated machine data bank when a broker "telephones" the computer. Vocabularies of from 1000 to 1500 words are the future objective of researchers.

The 'blue-skies' area. Way down the road is the concept of "pressless printing," in which holography will be employed. RCA's David Sarnoff Laboratories is engaged in such research. By using laser beams and the hologram technique, an entire page of *The New York Times*, for example, could be theoretically reduced and reproduced to a microfilm size, and a complete, 96-page issue of that newspaper could be contained on a very small microfiche.

These microimages could be magnified and projected at full newspaper size, page by page, onto an electrostatically charged web traveling at the incredible speed of 6700 meters per minute. The hologram image would be printed electrostatically, as in xerography, by a strobe exposure timing process onto the newsprint as it moves at this ultrahigh velocity. At this rate, a copy of *The Times* could be printed in a fraction of a second, and an entire edition could be run off at speeds that are comparable to the present high-speed press rate of 60 000 copies an hour.

A peek at Part II

Automation of the daily newspaper is less than half of the story of the dramatic revolution in the PPI. The final installment of this article will take up the electronic distribution and decentralized manufacture of magazines, periodicals, and books. The recent developments in press and binding equipment design are but a prelude to even more remarkable systems now in the R & D stages.

REFERENCES

1. Moore, J. K., and Cavanaugh, J. F., "A picture worth a thousand words," *Electronics*, vol. 40, pp. 113–121, Apr. 3, 1967.
2. Doebler, P. D., and Tetlow, J. S., "The shape of the future: knowledge—its assembly and organization," *Book Production Industry*, vol. 43, pp. 51–55, Jan. 1967.

High-field superconductor technology

Supermagnet technology is now at the stage at which the successful operation of comparatively few very large magnets has resulted in considerable optimism for future development

Charles Laverick   *Argonne National Laboratory*

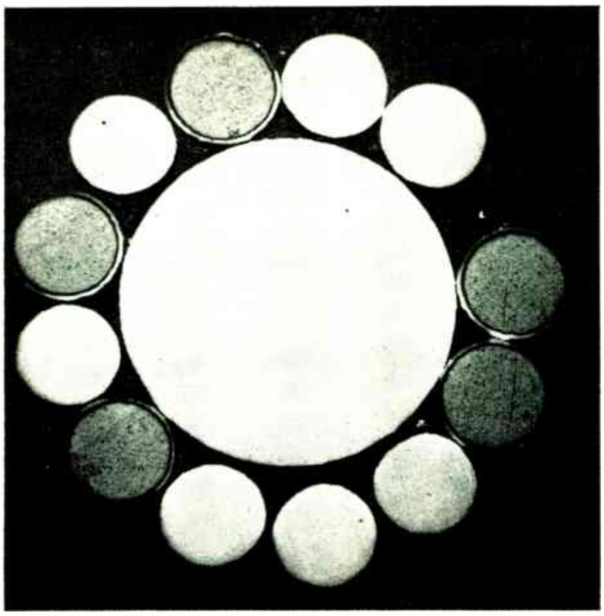
The past two years have seen the acceptance of superconducting magnets for engineering installations, the development of more advanced and robust engineering materials than the earlier conductors, and the wider application of high-field superconducting magnets in development programs for high-energy physics and thermonuclear fusion research. Their principal advantage is the great saving in operating cost when used in large-volume magnets up to 3 teslas. At higher fields a further advantage is their lower capital cost as compared with that of conventional electromagnets. They also have the potential, not as yet realized, of being more compact than a conventional electromagnet of identical performance.

The characteristics of stable magnets of low average current density (up to 1000 A/cm^2) are well understood, and such coils can be designed using precise engineering principles. Unstable magnets are not as yet amenable to exact design analysis, but there is now greater understanding of the mechanism of heat generation in superconductors and more experimental data on the performance of the various superconducting materials in coils. Both stable and unstable coils will operate at their design field for long periods of time if required. The principal difference between them is that a stable coil can operate with a section of the superconductor in the normally conducting state, whereas the normal region will spread throughout an unstable coil, which will then discharge. Stable coils will become unstable if the degree of overcurrent is excessive. Stable performance is obtained at some sacrifice in average current density as sufficient normal conductor is in intimate contact with the superconductor to permit the coil current to be shunted into it at the normal region without a significant increase in conductor temperature. Under this condition of operation the magnet is being operated as a cryogenic magnet and the normal design criteria for such magnets now apply.

Low-field magnets up to 4.9 meters (16 feet) ID at 1000 A/cm^2 are under construction. A 13.7-tesla high-field magnet with a 11.5-cm (4.5-inch) ID room-temperature bore has been operated at about 10 000 A/cm^2 and small-bore magnets up to 12 teslas are being offered as regular commercial items. Materials manufacturers can now supply conductors to meet a large range of design requirements on request and present-day cryogenic engineering technology is sufficiently advanced to meet the needs of most supermagnet designers.

Supermagnet types

It has become the practice to classify the magnets according to the degree of electrical stability they achieve in operation. The three main types of operation can be classified as stable, partially stable, and unstable, depending upon the operating current and upon the efficiency of the heat-transfer mechanism from the superconductor to the liquid-helium bath. Any superconducting magnet can act in any of these ways, depending upon the operating conditions chosen. In fact, even a cryogenic magnet will act as an unstable magnet and discharge itself if its operating characteristics are exceeded.



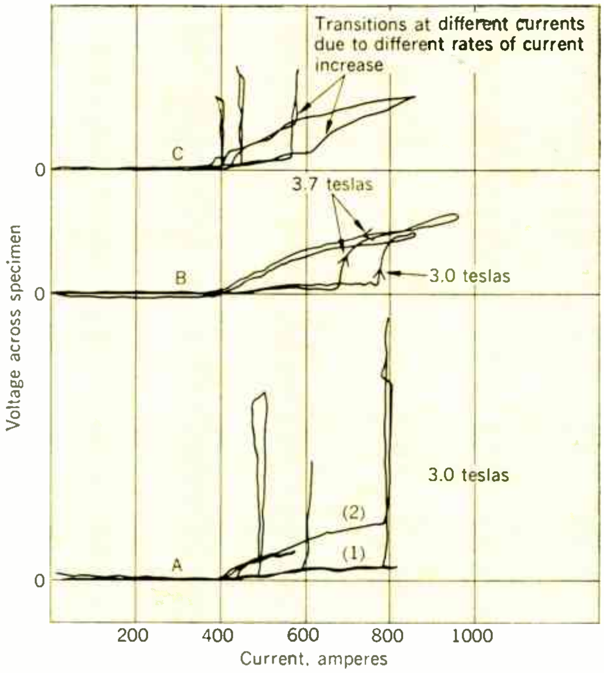
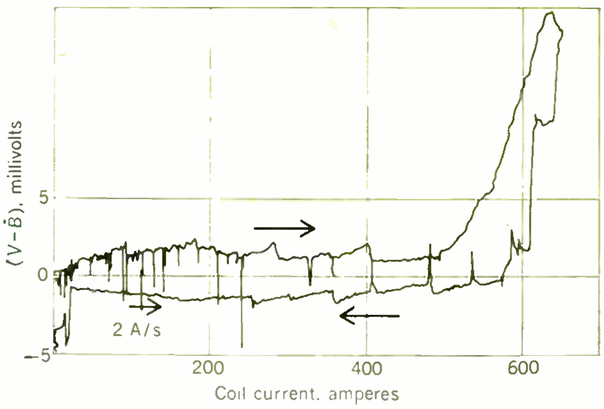


FIGURE 1. Short-sample characteristics for NbTi cable containing five 0.05-cm-OD copper-plated NbTi strands.

FIGURE 2. Variation of the difference signal ($V - \dot{B}$) with current for cabled NbTi coil shown in Fig. 1.



Stable operation is preferred for larger-size coils at the present time. Thus the very large bubble-chamber magnets now being planned or constructed operate in the fully stable mode. Stable performance is obtained by designing the superconducting magnet as a cryogenic magnet with a superconducting short circuit in the winding. If the operating characteristics of the superconductor are exceeded, the current will be shunted into the normal material and heat will be dissipated in that region of the normal conductor.

A cryogenic magnet is one that contains only normally conducting material (usually high-purity aluminum) and is cooled by some low-temperature liquid (such as hydrogen, which boils at 20.3°K at a pressure of 760 mm Hg). Successful operation of such a coil depends upon transferring the heat generated in the conductor by ohmic heating to the liquid bath through the conductor surface exposed to the liquid. Normal convection cooling in such a coil is effected through nucleate boiling of the liquid. As the surface temperature of the conductor increases, the boiling becomes more violent and film boiling begins. Since the heat transfer is almost one order of magnitude worse in this latter condition, the conductor temperature increases rapidly. Such magnets are usually cooled by forced circulation. A magnet operating in the stable region has a voltage–current characteristic that is reversible if the critical current of the superconductor is exceeded. Partially stable operation occurs when the zero-voltage point on current reduction occurs at a lower current than that at which resistance first appears when the current is increased. The operation is completely unstable if the coil discharges during current increase without any warning of the onset of resistance appearing. Examples of operation in the stable and partially stable regions are given in Figs. 1 and 2. These and the following figures lead to some understanding of the performance of practical superconducting coils.

In curve A of Fig. 1, several records are superimposed upon each other. The lowest record represents the terminal voltage–current record. The vertical voltage impulses at approximately 500, 600, and 800 amperes result from the application of a heater pulse using a heater coil around the sample. In this case, on recovery from the heater pulse, the voltage across the sample recovers to the higher line (2). As the current is decreased, the voltage across the sample (curve A) decreases until it returns to zero at 400 amperes. In this case, the information obtained from the heater test and the terminal voltage–current characteristics agree. The reduction in current (curve B) due to increased field can be seen. The transition currents (curve C) differ for rates of current increase of 2 A/s and 4 A/s, being higher at the highest rate of charge. These effects have been observed in short samples of ordinary copper, and are probably attributable to the fact that it takes a finite time for the copper to heat at any given current.

It can be seen that the normal regions detected in the short sample at approximately 400 amperes (Fig. 1) were not detected by the difference recorder in the coil test (Fig. 2). Such premature voltage increases can be caused by poor contacts; in this case, however, the contacts were good. The effect may, therefore, be caused by flux jumps of the type shown in Fig. 2, which are difficult to detect because of the small size of the sample. These jumps act in a similar manner to the heater pulses, generating heat in small local regions of the superconductor and causing limited current transfer in regions in which there is a poor bond resistance between the superconductor and the copper.

The flux jumps for the coil containing NbTi cable are much more frequent at lower values of current and occur much more often on current increase than on current decrease. The difference signal provides an accurate and useful indicator of the onset of normality for this heavily shunted cable. This can be seen in the region between 500 and 600 amperes, where the difference voltage rises markedly due to the onset of resistance in the coil.

The voltage across the coil decreases in steps, rather than smoothly, to approximately 580 amperes as the current is decreased from 640 amperes. The reason may be that the various superconducting strands are recovering at slightly different currents. Voltages across the coil have to be kept low to obtain these signals.

The voltage pulses at lower currents in Fig. 2 represent rapid movements of flux within the coil and flux jumps in the conductor. During these jumps the heat dissipated in the superconductor is conveyed to the helium bath through the normal conductor and the cooling channels. If less copper were used in the conductor, any one of these pulses might cause an instantaneous temperature rise within the conductor, which would lead to the discharge of the magnet at these lower currents. In this case, the magnet would be operating in the unstable region, a condition that was common to all of the early high-field magnets.

Some further insight into coil and conductor performance can be obtained by considering Figs. 3 and 4, in which a 13-strand cable containing nine NbTi strands has been tested by means of the heater test. The relevance of this test is well established by the similar disturbances usually observed in coils and recorded in the previous figures. Curve A in Fig. 3 gives the terminal characteristics for the cable at zero magnetic field. Both current increase and decrease are shown. It is seen that the two curves almost coincide. Normal regions exist in the cable

FIGURE 3. Short-sample performance characteristics for cable with nine NbTi strands.

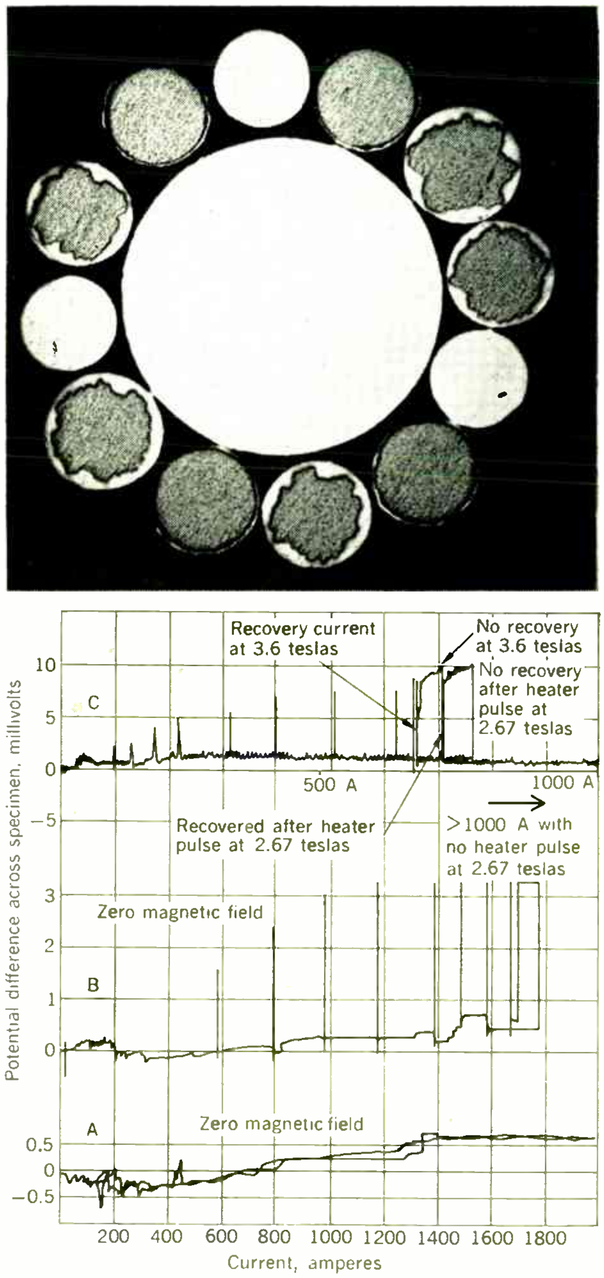
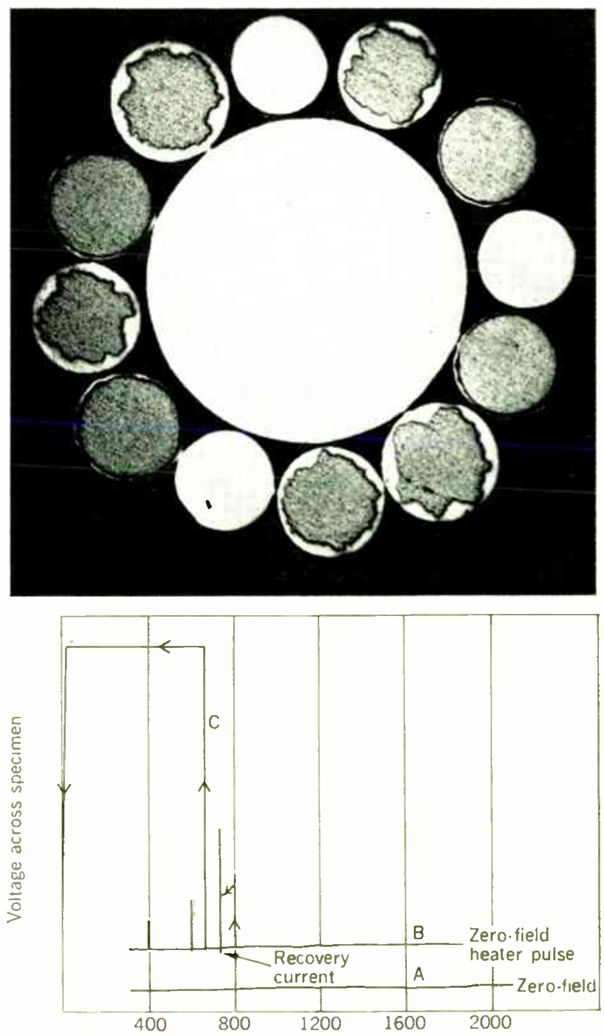


FIGURE 4. Short-sample performance characteristics showing propagation of a transition region below the recovery current for a large heat pulse.



and, at lower currents, there is some indication of negative-resistance effects, which are not understood. The dissipation in this sample at 2000 amperes is approximately one watt. The use of approximately equal heater pulses (curve B) for the first four pulses indicates the increasing sensitivity of the cable to disturbances as the current is increased, since the voltage generated in the sample increases markedly with each successive heater pulse. The heater pulse at 1800 amperes, with recovery on current decrease at 1700 amperes, gives the recovery current for this cable. In case C, the characteristics of the cable on an expanded scale (1000 amperes full scale) are shown. In this case, the terminal-voltage characteristic does not coincide with the recovery currents obtained from the heater tests. The decrease in recovery current at 3.6 teslas for this cable can be seen from curve C.

The need for care in applying heater pulses and interpreting the results of heater tests can be seen from Fig. 4, in which the recovery current at zero field is approximately 700 amperes. If a slightly larger heater pulse is now applied at some current below the recovery current, it can be seen that the specimen quenches completely and the transition propagates throughout the specimen. This is shown by the discharge curve C. It should be pointed out that an attempt is being made to derive more information from these heater-test records than was the case when the heater test was originated. This is being done here with the object of gaining further insight into coil and cable performance rather than merely predicting coil performance from short-sample data. The heater test was conceived for this latter purpose and has been very useful when employed in this way.

The parameters of most interest in coil testing are coil current I, magnetic field B, rate of change of magnetic field dB/dt or $\dot{B}$, voltage V across the coil, and the contact resistances between successive windings and at the junction of the incoming power leads. It is also convenient to have a device that permits measurement of $(V + a\dot{B})$ or $(V - a\dot{B})$, where a is a constant that can be adjusted to make $a\dot{B} = L\dot{I}$. Originally suggested by J. Purcell of the Argonne National Laboratory (ANL), it was thought that a unit that detected $(V - a\dot{B})$ would enable one to detect normal regions in coils easily, since $V = L\dot{I}$ (inductive voltage) + IR (resistive region) and $L\dot{I}$ could be adjusted to equal some multiplying constant times $\dot{B}$. Such a condition is difficult to obtain, particularly at low currents and high sensitivities, but this type of measurement is useful for monitoring coil performance. In many coils L changes considerably with current at low currents because of the shielding effect of the superconductor, and for this reason $(V - a\dot{B})$ changes at constant $\dot{I}$. The voltage V across the coil tends to remain constant in many stable coils when a transition occurs, since IR appears to increase at about the same rate that $L\dot{I}$ declines when a resistive region is created. This type of behavior is dependent upon the characteristics of the power supply used to charge the coil.

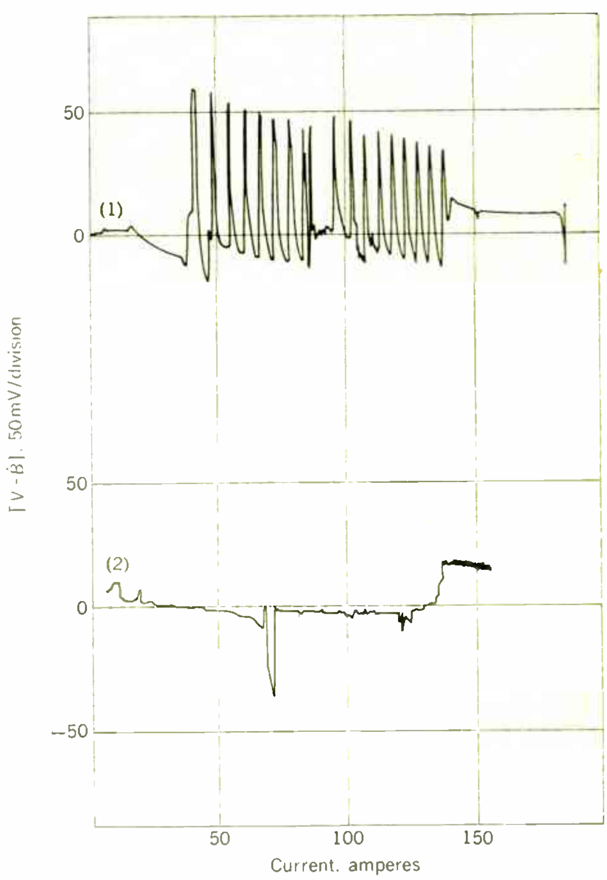
These points are clarified and the behavior of superconducting coils and cables better understood from a consideration of Figs. 1 to 6 inclusive. The characteristics of a coil wound from 2250 meters of an indium-impregnated cable containing seven 0.025-cm-diameter Nb25%Zr wires, copper coated to a radial thickness of 0.025 cm and wrapped with twelve 0.025-cm copper wires, shown in Fig., 5 should be contrasted with the characteristics of the coil (Fig. 2) wound from 590 meters of the NbTi cable (Fig. 1). The NbZr coil was well shunted with a low-resistance shunt to provide protection during discharge, whereas the NbTi coil was not protected in this way. Thus the NbZr coil had a much longer time constant for charge and discharge than the NbTi coil. In both cases, it can be seen that large excursions in $(V - a\dot{B})$ occur while the current is being increased slowly and linearly with time but that the disturbances are sharper for the NbTi coil, presumably because of its shorter time constant. The fact that the curve for current increase (curve 1) is different from the curve for current decrease (curve 2) indicates the manner in which the flux enters or leaves the bore when the current is changed.

The information derived by monitoring the parameters suggested can be contrasted by studying Fig. 6, in which the coil of Fig. 2 was energized in sequence a number of times and the variation of one important parameter with current was recorded each time.

The quantities studied were: (1) rate of change of magnetic flux $\dot{B}$ at the center of the inner bore of the coil, (2) voltage V across the coil, (3) the difference signal $V - a\dot{B}$, where a is a multiplying constant arranged to make $V - a\dot{B}$ approximately equal to zero at low values of current, and (4) $V + a\dot{B}$.

Comparing the curves during current increase, we note that a step in the difference signal (curve 3) occurs at approximately 170 amperes. This corresponds to a step in-

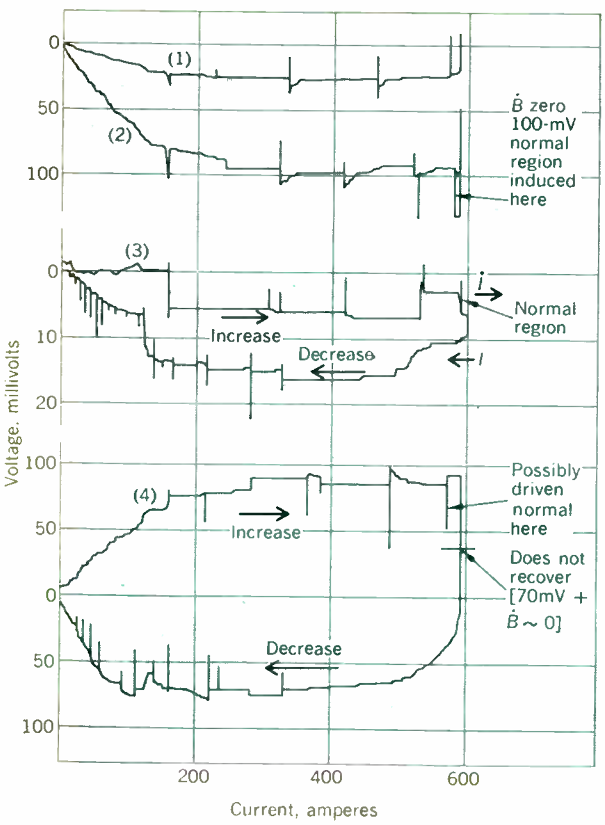
FIGURE 5. Operating characteristics for an Nb25%Zr coil; $(V - \dot{B})$ vs. current. (1) Current increasing and (2) current decreasing, at 0.5 A/s.



crease in the sum signal (curve 4), which is more difficult to detect, and to pulses in $\dot{B}$ (curve 1) and V (curve 2), which, in the absence of further information, might be regarded as flux jumps. This 6-millivolt step corresponds to the introduction of a normal region, which, at 500 amperes, would be the same as a power dissipation of 3 watts. A power dissipation of this magnitude would be unacceptable in many coils. A comparison of Figs. 2 and 6 reveals that the discontinuous flux changes during charge are much more frequent during the earlier charge (Fig. 2) than on later current increases (Fig. 6). The coil was never driven to a current at which it would discharge through self-propagation of a normal region, so the temperature changes in the coil were always small. The variation in behavior is probably due to progressive magnetization. Since the curves of Fig. 6 were not derived during the same charging sequence, it is not possible to look for a correspondence in all the disturbances recorded. The disturbances in any record corresponding to sudden flux changes often occur at different currents on subsequent current increases. Each flux jump during charge (curve 1) results in a small step increase in $\dot{B}$, although the rate of increase in current is constant. This jump must be due to a sudden increase in flux linkages at this current as the flux penetrates more of the turns in the winding.

The coil is finally driven normal at about 600 amperes, resulting in a pulse in $\dot{B}$, which then returns to zero when the current is held constant. The normal region appears as a step in V (curve 2), which consequently returns not

FIGURE 6. Variations of (1) $\dot{B}$, (2) coil voltage V, (3) $V - \dot{B}$, and (4) $V + \dot{B}$ with current for NbTi cabled coil, using cable shown in Fig. 1.



to zero but to a residual level of 50 mV as $\dot{B}$ is reduced to zero. At 600 amperes, this corresponds to a power dissipation of 30 watts, which is not acceptable. The sum voltage record also indicates the pulse at the onset of the normal region and the residual 50 mV at maximum current with $\dot{B} = 0$. The pulses during current decrease are again much more frequent in curves 3 and 4 than in the case of current increase. All current changes were set at approximately 2 A/s for these tests.

It seems necessary to monitor all four parameters continuously, and preferably simultaneously, while charging at low rates of current increases with low-noise power supplies in order to diagnose coil performance accurately. When the coil performance characteristics are understood, the coils can be operated with low-quality supplies if desired.

Magnet materials and conductors

The materials currently available include niobium zirconium, niobium titanium, and niobium tin, which are obtainable from a number of manufacturers in several countries for immediate use.

Information on the maximum current density at a given field with each material can readily be found in the extensive literature on the subject. Approximate upper critical fields and temperatures for the three most popular materials are NbZr: 6–7 teslas, $T_c = 10°K$; NbTi: 12 teslas, $T_c = 9°K$; and Nb_3Sn: 25 teslas, $T_c = 18°K$. Typical current densities are as follows: for NbZr, 4×10^5 A/cm² at 2 teslas; for NbTi, 10^5 A/cm² at 4 teslas and 4×10^4 A/cm² at 8 teslas; for Nb_3Sn, 10^6 A/cm² at 2 teslas and 10^5 A/cm² at 10 teslas.

The current densities attainable in bulk materials are greater than those obtained in composite conductors, since the former contain normal conductor as well as superconductor. Conductor characteristics vary with the nature and distribution of the defect structures in the superconductor. Short-sample current-carrying capacity is enhanced and the inherent stability of the conductor is increased as the conductor dimensions are reduced. Conductors with the most favorable surface-to-volume ratios are more stable because the conductor cooling is best for such arrangements. Superconductors with the highest critical temperatures give the greatest working temperature margin and are to be preferred if all other considerations are equal. Choice of conductor type is now dictated largely by price and by attainable average current density in the windings.

A promising new development in practical high-field superconductors is the technique developed by Tachikawa of the Institute of Metals, Tokyo, for fabrication of V_3Ga in flexible wire or strip form. This can also be made as a composite with copper. The short-sample characteristic obtained is approximately that of present-day commercially manufactured Nb_3Sn at magnetic fields of up to 15 teslas and is much higher beyond this field. The U.S.S.R. and Japan are manufacturing three-element systems of NbZrTi, which are comparable in properties with currently available NbTi in the United States, but such systems do not, as yet, appear to offer any significant advantages over NbTi. Niobium tin and niobium zirconium are also in use in Communist China.

At present, magnets are being fabricated from NbTi, Nb_3Sn, or both, usually as composite conductors in intimate contact with copper, and frequently with the addition of reinforcing at high operating stress levels. The properties of these basic superconductors are well known. NbZr is now obsolescent, and current practice favors the use of the cheaper, lighter, and more ductile NbTi. Several recent, highly successful large coils have been fabricated from stranded superconductors, but improvements in the manufacture of composites of copper-clad NbTi with one or more superconducting strands and almost any desired current-carrying capacity or size render such stranded conductors obsolescent.

Experience with cables and composites has shown that metallurgical-type diffusion bonds between the superconductors and coextruded copper give best performance and are mandatory when superconductors of large cross sections are used. Mechanical-type bonds are being used with success in composites containing superconductors having small cross sections. When the highest current density is desired, the diffusion-type bond permits the use of much less copper in the cross section of the composite for the same performance, since the contact between the two metals has minimum resistance to the passage of electricity.

The techniques of adding extra copper and reinforcement are also being applied to the 1.25-cm-wide Nb_3Sn ribbons now available; thus, Nb_3Sn conductors of almost any strength and current-carrying capacity can also be fabricated. In this case, soldering techniques are used to add the extra copper.

Nb_3Sn is very brittle and can be operated only in compression. The superconductor in the RCA ribbon is compressed in manufacture by differential contraction of the Nb_3Sn layer and Hastalloy substrate on cooldown after the vapor deposition process. The General Electric Company has developed a range of composite conductors containing Nb_3Sn diffusion layers on a niobium substrate to give the current desired, copper cladding to any thickness to give the specified resistance per unit length, and stainless-steel cladding to give the strength required. The combination of copper and stainless steel with the superconductor in the center puts the superconductor under high compression after cooldown from the soldering temperature. Thus, when the conductor is strained, the superconductor is operated only to the limit of zero compression and never into the positive stress region. Naturally, similar considerations can apply to conductors of niobium zirconium and niobium titanium.

The most advanced heavy section conductor yet produced has been supplied to Argonne National Laboratory by the Supercon Division of National Research Corporation for use in the 4.9-meter-ID 2-tesla magnet of the Argonne 3.6-meter-diameter hydrogen bubble chamber, the first of a new generation of giant particle detectors. The strip contains sufficient copper to support the hoop stresses that are generated during magnet operation, has a width of about 5 cm and a thickness of 0.25 cm, and contains six strips of niobium titanium capable of carrying 3000 amperes at 3 teslas. The strip is anisotropic, and conducts a higher current when the magnetic field is parallel to the large dimension of the NbTi conductor cross section than when the field is perpendicular to it. A more advanced conductor also has been produced by Supercon for the Brookhaven National Laboratory half-scale model of its proposed 4.2-meter-ID hydrogen bubble chamber. Specifications for the two conductors are given in Table I.

I. Comparison of superconductor specifications

	Argonne Magnets	Brookhaven Magnets
Width, cm	5.0	5.0
Thickness, cm	0.25	0.20
Piece length, meters	220	373
Piece weight, kg	245	330
Critical current,* amperes	3000 at 2.5 T (P) at 1.35 T (N)	5880 at 2.7 T (P) at 2.05 T (N)
Resistance ratio (300°K/4.2°K)	200	150
Design stress, N/cm²	8300	11 000

* Fields parallel (P) and normal (N) to the long side of the rectangular sectioned strip.

The most advanced NbTi composite conductors for use in high-current-density smaller coils of up to 60 cm ID at central fields of about 6 teslas are of the type being produced for the outer sections of the 51-cm-ID, 8.8-tesla coils being constructed for the NASA Lewis Research Center by the Avco-Everett Research Laboratory. The conductors can be wound into stable magnets if they are suitably cooled. They contain a considerably lower ratio of copper to superconductor than earlier conductors of this type. The increased stability is attained by limiting the temperature rise in the superconductor when sudden movements of flux occur. This is accomplished by using a larger number of superconducting strands for a given total cross section of superconductor so as to maximize the surface-to-volume ratio for the superconductor and, hence, obtain the best possible cooling. Extra superconductor is also being incorporated in the conductors where necessary as a simple and cheap means of providing additional reinforcing in highly stressed conductors.

Proposals have been made for hollow copper-clad conductors in which high-pressure helium in the dense fluid phase can be circulated so that high heat-transfer rates can be obtained. Tests on samples of this type of conductor have been made at the Stanford Linear Accelerator Center, and it is suggested that this scheme will allow high current densities to be obtained in larger coils without the necessity for a liquid helium tank encasing the complete coil.

Present Nb_3Sn composite conductors are both more rugged than formerly, more stable in operation, and have a higher current-carrying capacity. The most advanced conductor of this type was recently developed by General Electric to meet the Avco-Everett requirement for the inner sections of the previously mentioned 51-cm-ID, 8.8-tesla magnet. The 1.3-cm-wide conductor contains outer layers of 0.0076-cm-thick stainless steel for reinforcing and the superconductor is sandwiched between two layers of 0.005-cm-thick copper to provide the electrical protection. The reinforcing has been provided in order to meet a hoop stress requirement, but work at Argonne has shown that this extra strength is also desirable to give a conductor that is easily handled in use, will support high edge-to-edge loads without damage to the ribbon, and will provide liquid-helium cooling on the face of the ribbon when this is wound into a pancake-type coil.

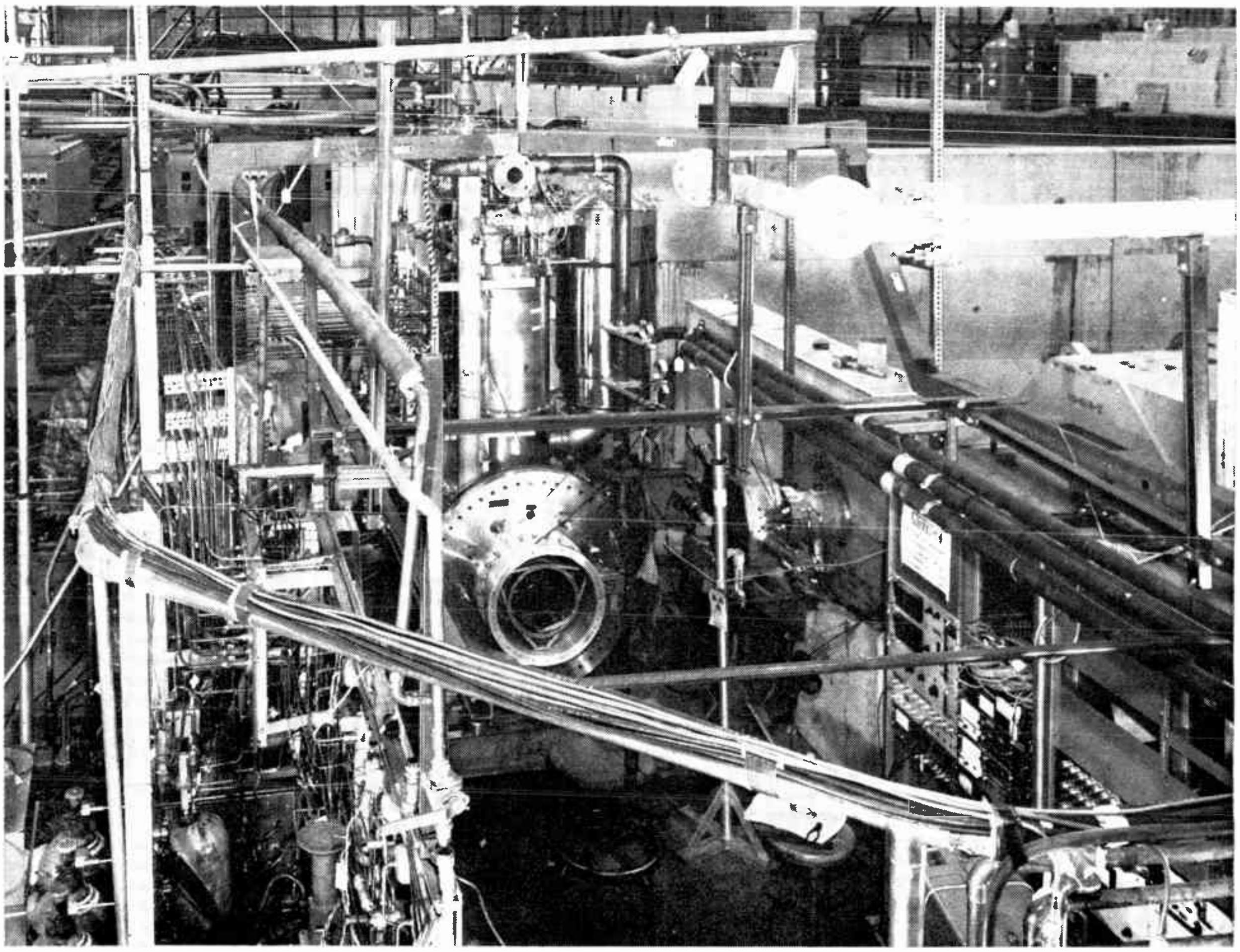


FIGURE 7. General view of the 25-cm superconducting helium bubble chamber in meson experimental area of the Argonne 12.5-GeV zero-gradient synchrotron complex.

The conductor is rated at 300 amperes for 10 teslas. Recent ANL tests with 18-cm-ID counterwound pancakes of this material show it to be stable in operation at low fields up to 900 amperes with a smooth reversible transition into the resistive region. Experiments at Argonne with larger-size pancake-type coils also show that higher operational stability in 1.3-cm-wide copper-coated Nb_3Sn strips can be obtained if a number of narrower parallel superconducting strips are used instead of one superconducting strip that occupies the entire width of the superconductor.

It is still true at the moment that NbTi composite conductors provide the cheapest means of providing fields of from 6 to 8 teslas in medium-size coils of about 20 cm ID or larger and that Nb_3Sn composite conductors are necessary for the higher fields. The techniques for using Nb_3Sn strips in pancake-type windings to provide the highest fields and highest current densities in coils larger than 5 to 10 cm ID are still in their infancy and require considerable development. A great deal has to be done before the potential performance of present materials, as measured by the attainable current densities and maximum attainable magnet fields reached in short samples, can be realized in practical magnets.

Present magnets

The Argonne 28-cm-ID, 4.5-tesla superconducting helium bubble-chamber magnet, which consists of the two outer radial sections of the 6.7-tesla coil, is still the only superconducting magnet of any appreciable size (64 cm OD, 30 cm long, stored energy $\approx$ 500 000 joules) to have been operated in a complex engineering installation. This has generated a great degree of confidence in such magnets. At Livermore, a minimum-B baseball magnet of 25-cm diameter also has been successfully installed for use in plasma-physics experiments. It was natural in these early magnets to concentrate on the magnet problem but it is now obvious that the emphasis should be on the complete system design with perhaps more emphasis on the cryogenic problems than on the magnet. The system comprises the cryogenic environment, the conductor and coils, a mechanical support structure, a protective system of some kind, and a means of energizing the coils. The complexity of the final bubble-chamber system, as operated in the experimental areas of the ANL 12.5-GeV synchrotron complex, can be seen from Fig. 7.

One large 3.7-tesla, 30-cm-bore, 3.6-MJ transverse field coil with a uniform transverse field region of the order of 1.5 meters long has also been successfully operated, but no attempt is to be made to incorporate this in a complete system.

RCA has successfully built and operated a 15-cm-ID, 13.7-tesla magnet having an 11.5-cm (room temperature) bore. Like the 25-cm bubble-chamber magnet system, this magnet is unstable and has been safely discharged at maximum field. The 2 MJ of stored energy is absorbed by copper secondaries distributed throughout the coil to ensure a reasonably uniform temperature distribution during discharge.

As mentioned previously, cables have been rendered obsolescent by advances in manufacturing techniques. However, several cabled-type magnets wound from these copper-cladded stranded superconductors and stabilized by the addition of extra copper have been successfully designed and operated at magnetic fields up to 7 teslas in bores of the order of 30 cm in solenoid-type configurations. Some of these magnets have also been incorporated into experimental apparatus where they have continued to perform successfully. Average current densities for such coils have been in the range of 4000 to 5000 A/cm^2. Consequently there is no doubt that the use of the more advanced conductors in coils of this nature will lead to even higher current densities.

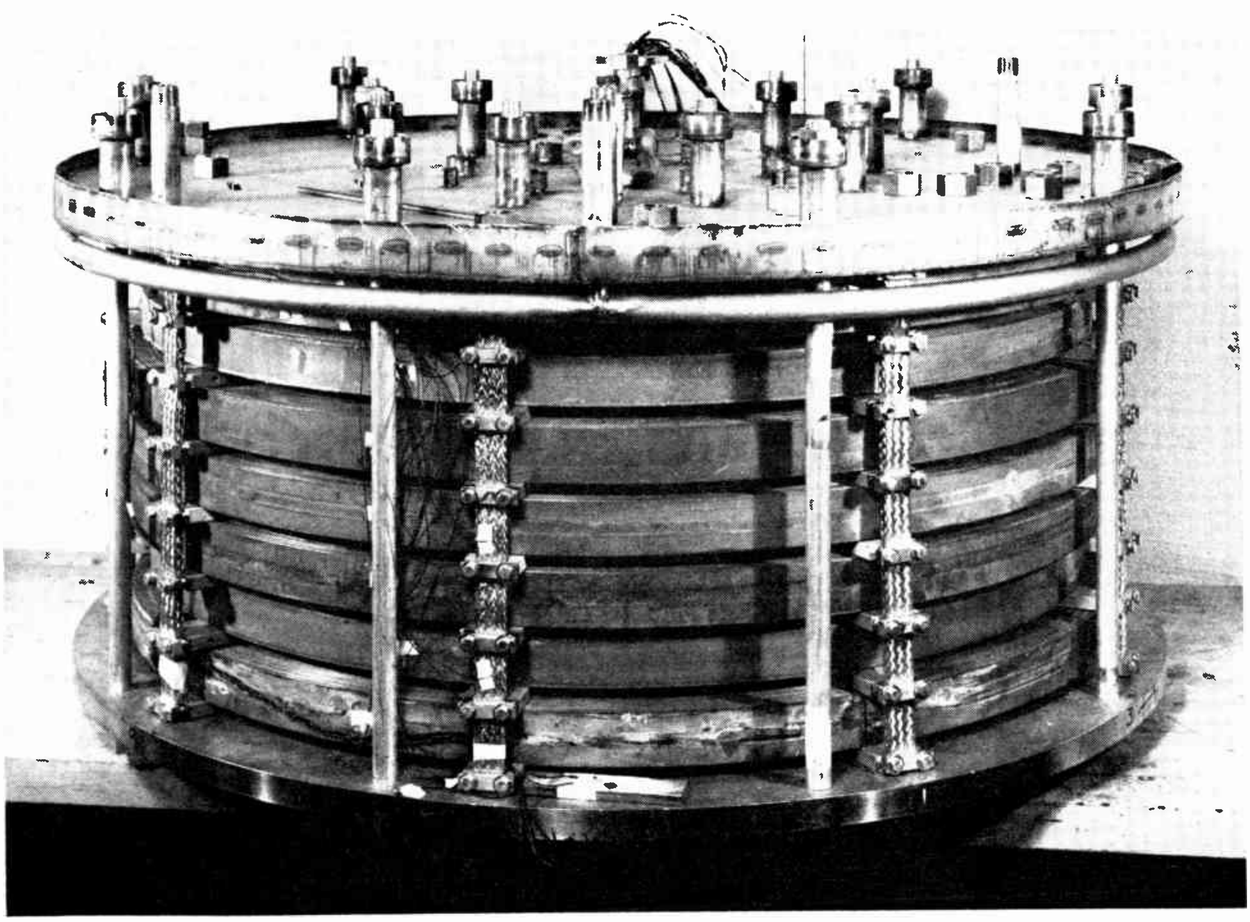
The first three lengths of strip delivered for the Argonne 5-meter-ID magnet were each cut in half and wound to form the six-layer coil shown in Fig. 8. This coil has a 61-cm ID and an 86-cm OD, stands about 38 cm high, and weighs about 800 kg. The total winding length is 700 meters. Teflon tape and epoxy form interturn insulation.

When tested, this coil carried 4600 amperes while developing maximum fields of 3 teslas parallel and 1.9 teslas normal. The 1.9-tesla normal field was the critical condition and corresponded to the actual short-sample data that had been measured previously in the strip. Current sharing between the copper and superconductor commenced at 4650 amperes, thus demonstrating the good thermal and electrical properties of the bond. This sample of strip exceeds the minimum specifications of Table II, an ample reservoir of performance stability.

The implications of this test magnet for the construction of low-current-density magnets of this type are best understood by considering the simplicity and strength of the coil (see Fig. 8). The main specifications of the Argonne and Brookhaven National Laboratory magnets as now being constructed are given in Table II. One pancake of the Argonne coil is shown in Fig. 9 on its winding jig to give an idea of the scale of the work.

Some of the philosophy in design and procurement which has developed during the ANL project represents a new trend—one that points the way to the future. Prior to this project, designers tended to accept commercially available materials and use them in the best possible way in a given magnet design. This project is sufficiently important that the designers could afford to design an acceptable conductor and specify its properties to the manufacturer. CERN in Geneva has followed this approach

FIGURE 8. Advanced 0.6-meter-ID pancake-type coil of simple, rugged construction using the conductor designed for the 4.9-meter-ID Argonne bubble-chamber magnet.



II. Comparison of bubble-chamber magnets

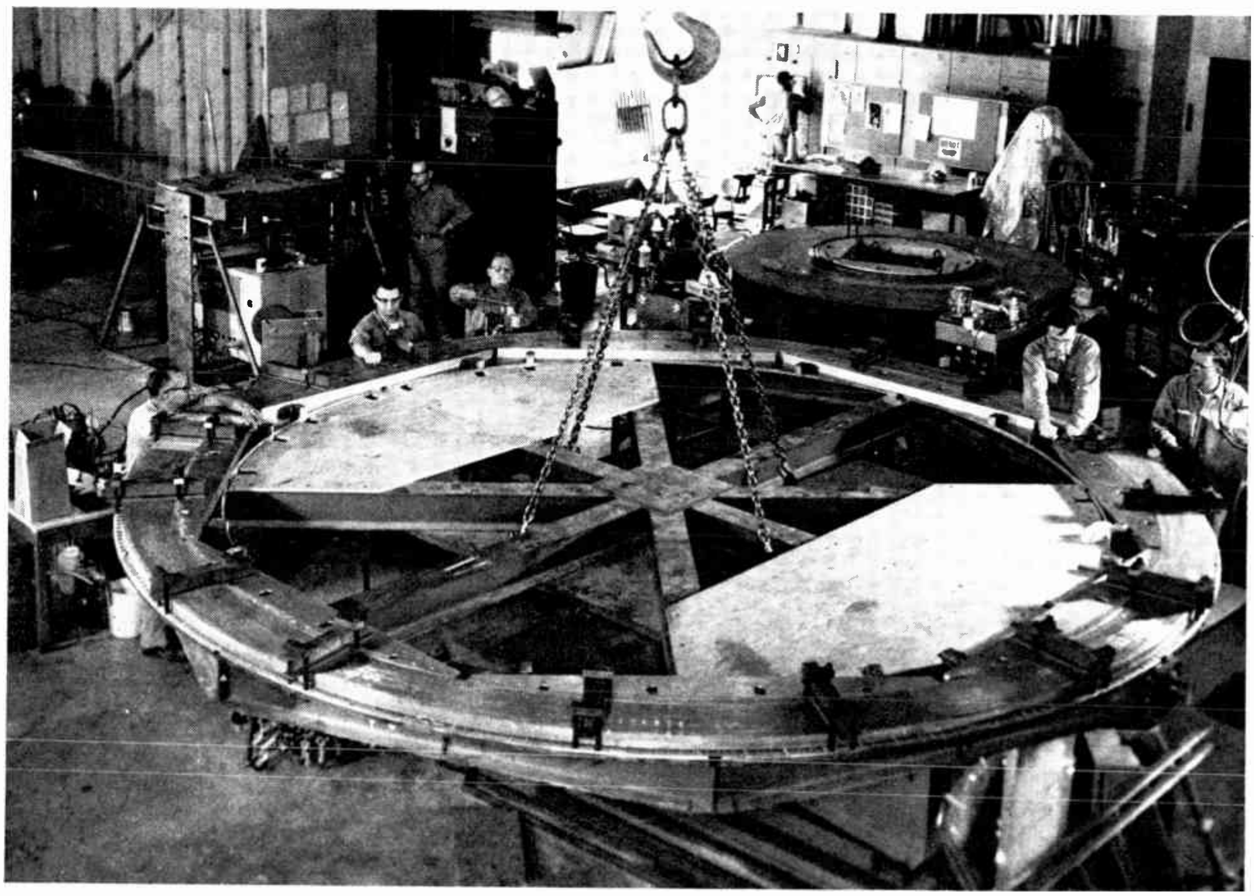
	Argonne Magnet	Brookhaven Magnet
Inside diameter, meters	4.83	2.44
Outside diameter, meters	5.33	2.74
Length, meters	3.1	2.25
Central field, teslas	2.0	2.5
Field energy, megajoules	80	45
Design current, amperes	2000	4500
Conductor length, meters	41 600	11 800
Conductor weight, tonnes	45	11
Conductor cost (less copper)	$460 000	$178 000

in specifying conductor for its new large bubble-chamber magnet. Samples were then fabricated and supplied by the various interested suppliers and subjected to suitable tests to determine which sample would be best for this coil. These included physical examination of the specimens, X radiography to determine the positions of the various conductors and the presence or absence of any flaws, photomicrography of the conductor cross sections, measurements of the electrical quality of the bonds between superconductor and copper, and, finally, measurements of the short-sample characteristics of the conductors at 4.2°K in the specified magnetic field, up to and beyond the maximum stress specifications for the conductor.

The technical requirements specified for the conductor and magnet included maximum field; maximum permissible heat leak due to power leads; maximum allowable power dissipation during charge, constant operation, and operation at maximum current; allowable percentage of overcurrent without self-propagation of a normal region, except if the liquid helium is lost; conductor current and size; copper resistivity; and maximum operating stress level. These requirements are typical of those that must be borne in mind by any supermagnet designer.

In the case of the proposed larger Brookhaven National Laboratory system, the potential saving in operating cost is almost four times more than that for the Argonne coil, since it would require almost 40 MW of electric power to generate 3 teslas by use of conventional copper coils. In addition to this saving in operating cost, the capital cost of the proposed system is now much less than that of a conventional system. Moreover, it weighs much less, so it is feasible to consider moving it from one particle beam

FIGURE 9. Single pancake of the 4.9-meter-ID Argonne hydrogen bubble-chamber magnet in position on its winding jig after completion of the winding process. Thirty such pancakes are assembled in the final 80-MJ, 2-tesla coil, which has a 1350-tonne iron yoke.



to another. The estimated cost of $16 million for this bubble chamber, including $3 million for buildings and experimental facilities and $3 million for contingencies, is approximately the same as the cost of the Argonne system bubble chamber.

The stresses in the conductor for the large Brookhaven system are much greater than in the case of the Argonne system, and it is now necessary to use reinforcing material in the winding. At the present time, reinforcing strips of either stainless steel or beryllium copper wound with the composite conductor have been considered.

The CERN group has planned a 3.6-meter hydrogen bubble-chamber system with a 3.5-tesla central field located in an iron house to provide magnetic shielding for the auxiliary experiment, and a preliminary design study has been completed. High-field chambers are more in favor at SLAC and the Rutherford Laboratory, and 1.5-meter-diameter 7-tesla systems are planned.

A novel high-energy physics experiment scheduled to begin soon is the Alvarez balloon-borne experiment, which combines the present-day sophistication of high-energy physics instrumentation with the older cosmic-ray techniques in an attempt to obtain information in the range of 300 GeV and beyond. The detectors will be located in a 1.5-tesla transverse magnetic field having a length of approximately 1.8 meters and a bore of 0.9 meter. The conductors are cables containing NbTi and the winding is of the saddle-shaped type necessary to develop a transverse field. The magnet was designed by C. E. Taylor of the Lawrence Radiation Laboratory, Livermore, Calif., and the coil was successfully tested in late 1967.

The development program in thermonuclear fusion is now deriving considerable benefit from these advances. A neutral injection experiment at Livermore will use a 0.9-meter-diameter baseball with a mirror ratio of 2:1 and a central field of 2 teslas for confining the plasma in a "minimum B" configuration. A new principle for confining plasma is to use toroidal magnetic fields generated by levitated superconducting rings. Such a system has been successfully operated by a Princeton group.

Bending magnets of aperture up to 15 or 25 cm with lengths up to 2 meters and field strengths up to 6 teslas are either planned or in construction in many countries concerned with the handling of beams of high-energy charged particles.

Cryogenic requirements

One of the most severe problems with very large systems is the problem of arranging a cooling-down system that minimizes thermal stresses, particularly those due to uneven temperature distributions during cooldown or discharging of the magnet system. In the case of the Argonne bubble chamber, the weight of the coil conductor is approximately 45 tonnes, almost all of which is copper. A cooldown time of about 60 hours is required to cool the conductor and helium vessel. Temperature stratification in the helium container is prevented by introducing a mixture of gaseous and liquid helium at the bottom of the tank under the lowest coil pancake. The rising helium provides sufficient stirring action to maintain a uniform temperature distribution. The steady-state refrigeration requirement is estimated to be less than 50 watts, but a reasonable cooldown time can be accomplished only with a larger refrigerator. For this reason, and to provide a margin of safety in operation, the refrigeration capacity has been set at 400 watts at 4.2°K. An additional 500 watts of refrigeration at 50–70°K is provided for thermal shield cooling. In the case of the Brookhaven system, a 1000-watt refrigeration capacity at 4.2°K is being provided, with liquid-nitrogen precooling.

Preliminary estimates of the refrigeration requirements have been made assuming that superconducting magnets would be used in the experimental areas of the new 200-GeV proton synchrotron to be constructed in Weston, Ill. The problems here differ considerably from those of the specialized large bubble-chamber coils, which are few in number. The experimental hall for this new machine is about 600 meters long and 120 meters wide. Before the relative advantages and disadvantages of superconducting and conventional systems can be compared, practical schemes have to be devised for each. Three main possibilities are: to provide a central liquefier system with a 600-meter-long liquid-helium distribution system; to provide separate liquefiers at each magnet system with a common compressor unit for each 10 or 20 magnets; or to provide separate liquefiers at each magnet station. In the first study, it was assumed that 164 magnets of the order of 2 meters long, 20 cm ID, and about 2 megajoules of stored energy would be required, of which 80 percent (130 magnets) would be in use at any one time.

The cheapest system considered envisaged a central liquefier with Dewar distribution at $7.1 million for a total of ten-year operating cost plus capital cost. A system with individual magnet refrigeration was estimated at $13.5 million, with the refrigeration units priced at $31 000 each. This is, of course, an initial estimate and will provoke thought and discussion. It has already produced a swift reaction in one cryogenics company who feels that it could produce the estimated $13.5-million system for $8.3 million, using higher-capacity refrigeration units than those previously considered, priced at only $15 000 each with an additional $5000 per unit to represent the price of a central compressor for each 20 refrigeration units. Perhaps this type of reduction is indicative of things to come.

Present-day large superconducting magnets with resistive contacts and the potential for operating as cryogenic magnets have ohmic losses that have not been considered in the study. The variations in cost and power dissipation for the examples studied reveal that the cost of each extra watt (per magnet) of power dissipation for the total system is about $1.2 million per watt. This is based on the estimated capital cost plus ten-year operating cost for the total system. At these costs, it is worthwhile spending some time developing superconducting coils in the 2-MJ range with minimal power dissipation and highest possible current density. There would be little comfort in operating these coils as cryogenic systems in the stable mode at these prices in view of the proven reliability of systems that, in fact, will discharge if the critical current is exceeded.

Thus, in beam transport systems, the magnets must be designed for minimum internal dissipation, including lead losses, dissipation at contacts, and dissipation due to normal regions in the windings. They must also be designed to keep heat leaks to a minimum. The large magnets are special cases, but it would seem wise to limit internal dissipation due to contacts and normal regions to 5 or 10 percent of the static heat leak.

If the maximum estimated static dissipation for the ANL 3.6-meter bubble-chamber magnet is assumed to be 50 watts, a maximum permissible internal dissipation of 5 watts total for contacts and normal regions would result. The total number of connections in the coil will be approximately 100. If all the 5 watts is dissipated at the contacts, this gives an average of 50 mW per contact, corresponding to a voltage drop of 25 μV per contact at 2000 amperes. In practice, one or two contacts may begin to deteriorate, and thus cause excessive dissipation in the system. Tolerable limiting contact resistances should be monitored; for the present example, they are of the order of 12×10^{-9} ohm per contact, with acceptable voltage drops of about 25 μV per contact. The pressure to reduce heat losses due to input power leads and to reduce contact resistances will therefore increase rapidly. Reasonable operating losses for contacts and leads should be about 2 watts, assuming a dissipation of 50 watts including the heat leak. Operating currents of 2000 and 10 000 amperes are planned for such systems, which gives upper limits for contact resistance, heat leak through the leads, and tolerable power dissipation in the leads.

Beam-transport magnets and bending magnets will be designed to operate at maximum heat leaks of 2 watts to keep capital and operating costs low when large numbers are required for a given installation. Since the lead losses and contact losses should be less than the heat leak, maximum desirable operating levels should be about 0.5 watt for each parameter. Because heat leak is proportional to surface area, minimizing heat leaks implies maximizing average current densities in coils. Typical lead losses at 2000 amperes are one watt per lead with counterflow gas cooling. This may not be satisfactory for closed-cycle systems, since the gas is raised to room temperature rather than recirculated at a temperature of 10 or 20°K.

Power supplies with very low ripple content and precise electronic control of current increase and decrease are desirable for test purposes. Measured voltages in monitoring $\dot{B}$, I, B, V, $(V - a\dot{B})$, and $(V + a\dot{B})$ range from hundreds of microvolts to about 100 millivolts. Contact resistance of the order of 10^{-9} ohm are desirable at 2000-ampere operation, with correspondingly low values for 10 000-ampere operation. Charge times of the order of 10^3 to 10^4 seconds are desirable for testing coils in the 1- or 2-MJ range. If a 2-MJ coil is charged in 10^4 seconds, the rate of charge is 200 joules per second. A flux pump of 95 percent efficiency would dissipate 10 watts for the charge period, and one with 99 percent efficiency would dissipate 2 watts. At 2000 amperes, a total contact resistance and coil dissipation of one watt implies a voltage drop of 500 μV. At 10 000 amperes, this dissipation level implies a voltage drop of 100 μV.

Two modes of operation are promising. Coils of high stored energy can be charged by an external power supply, and a flux pump can be used to maintain a constant field. Present-day flux pumps are probably adequate for this purpose without further development, since a 2000-ampere output at 1 mV seems adequate. Such a device is self-limiting for many unstable magnets, since they will usually not discharge if such low voltages appear at some point in the winding. Coils in the 2-MJ range or less can be energized if flux pumps with 200 watts output at 99 percent or more efficiency, with output voltages of 200 to 300 mV can be made available. A 2-MJ coil at 2000 amperes has a one-henry inductance and can be charged in 10^4 second at 200 mV.

External dump resistors are necessary in larger systems to remove most of the coil energy on discharge. In the Argonne system, an external "dump resistor," which is always in the electrical circuit, is connected in parallel with the coil. If a resistive region develops in the coil, both the stabilizer and power-supply resistors will be disconnected from the circuit and the coil current will flow through the dump resistor. The energy stored in the magnetic field will then be dissipated in the dump resistor at an initial rate of approximately 2 MW (2000 amperes at 1000 volts). In the worst case, with the whole coil resistive, at least 90 percent of the energy will be dissipated in the dump resistor.

A homopolar generator could be used to charge a coil. Discharge could then be effected by switching the generator into the motoring mode with the coil as supply so that the electric energy can be fed back into the electric power system. Dump resistors at low temperatures could be considered. As experience is gained with the magnets, such protective systems may become unnecessary.

Supermagnet applications

The greatest potential applications at the moment are in high-energy physics, with its associated accelerator facilities. This category includes new bubble-chamber coils, beam-transport elements, and possibly spark-chamber coils. The possibilities for accelerators should be assessed separately. The magnets have application for some industrial consumers. Familiarity with the coils and the associated cryogenic techniques on the part of a larger number of people may lead to many novel applications that have not yet been foreseen.

The main possibilities for large-scale applications would appear to be MHD and fusion generators. It may be many years before successful prototypes are evolved. The fusion research studies involving steady fields are already committed to the use of superconducting magnets. This is a limited and small-scale use. Magnets are required in various fields of physics, such as solid-state physics; consequently, many superconducting magnets will continue to be required by university and industrial laboratories.

Some thoughts for the future

Materials. Although Nb_3Sn, NbTi, and NbZr have been exploited to a considerable extent, intensive studies of other possibilities (such as V_3Ga, V_3Si, Nb_3Al, and the niobium nitrides) and the search for new materials with higher critical fields or temperatures will undoubtedly continue.

There remains considerable incentive to develop materials that are either inherently stable or at least more stable than those currently available at comparable current densities. Programs to develop techniques for controlling amplitude, and distribution and density of flux pinning sites are typical of those that should lead to improvements in this area.

Conductors. Composite conductors of enhanced current density and stability comparable to that of conductors now available can be developed by utilizing the more stable materials mentioned earlier. Improvements in conductor resistivity, volume-to-surface ratio of included superconductor, and the interface bond resistance between superconductor and normal metal are also desirable. The use of aluminum, with its saving in weight in addition to its superior conductivity at higher magnetic fields, should also be encouraged. Improved techniques for bonding aluminum to various superconductors should be developed.

It seems unwise to develop 1200-ampere, 10-tesla Nb_3Sn ribbons with dimensions of the order of 1.3 cm wide by 0.015 cm thick and then lose a great deal of the advantage by bonding them to copper whose resistance at 10 teslas is three or four times that at zero field. Resistance ratios of about 1000:1 are easily attainable with aluminum.

The bond resistance between electroplated copper and NbTi wires is usually unsatisfactory. Metallurgical bonds are now being attained in some cladding processes and during the manufacture of Nb_3Sn tapes by the diffusion process. Tests to measure bond resistance are already being used, but it is necessary to establish satisfactory bond resistance limits for various conductors.

Higher-field magnets of the order of 30 cm ID at 10 teslas and medium-field magnets of the order of 3 to 6 meters ID at 3 to 4 teslas will necessitate the development of high-strength composite conductors.

Ultrahigh-current conductors and very-low-current conductors will become available. Techniques for fabricating and using conductors of much smaller dimensions than present conductors require development. As familiarity with operating techniques increases and detecting devices and safety devices are improved, the degree of overcurrent specified as a safety precaution in larger magnets will decrease and increases in operating current density will follow.

Operating experience with magnets of stored energy in the range of 1 to 5 MJ will undoubtedly lead to the adoption of safe, unstable designs operated within their stable limit so as to reduce coil volume and cooling area to a minimum while giving maximum current density. It is easy to see that this goal is not unrealistic.

Composite conductors that resemble normal large conductors, with a central space for the cooling fluid, have also been suggested because of the attraction of eliminating the need for helium tanks around the winding.

Improvements in diagnostic techniques will be developed. There is a need to ascertain and explain the operating characteristics of larger coils at various temperatures, with various degrees of shunting and wound from materials of different types and cross sections. Present coil studies, such as those reported here, are preliminary in nature and have only indicated where problems may exist. One possibility for damping out flux jumps would be to use longer-time-constant coils with eddy-current damping rings of high-purity aluminum distributed through the windings. Nb_3Sn ribbon coils have been found to carry high currents below the λ point at high fields. In NbTi and NbZr coils, both increases in the current-carrying capacity of some coils and reductions in the current-carrying capacity of others have been observed below the λ point. The difference is more likely to be caused by differences in surface-to-volume ratio for the various cases and by differences in the quality of the electrical bond between superconductor and normal conductor rather than to fundamental differences in the behavior of various types of superconductor. Degradation effects are probably enhanced at high conductor currents with poor bonding between normal conductor and superconductor. These various problems will receive study and should be resolved in the near future.

Contacts. Contact resistances of the order of one microhm have been used in many coils. Ribbon-type contacts of the order of 10^{-8} ohm are also easily attainable. Reductions in these values are essential as the need to keep the total system dissipation as low as possible grows. Truly superconducting contacts of various types, which are easy to fabricate, should be developed. Operation of many magnets in the persistent mode will become more and more common as confidence in this type of operation is gained.

Cryogenics. The cryogenic system is a natural part of the superconducting-magnet system. Cryogenic engineering technology is well established and the use of the best cryogenic techniques in present magnet systems has been limited probably only because of the novelty of the magnets themselves. Increasing familiarity with these techniques will develop as more people become involved in low-temperature work and as management becomes ready to authorize reasonable expenditures for this type of effort.

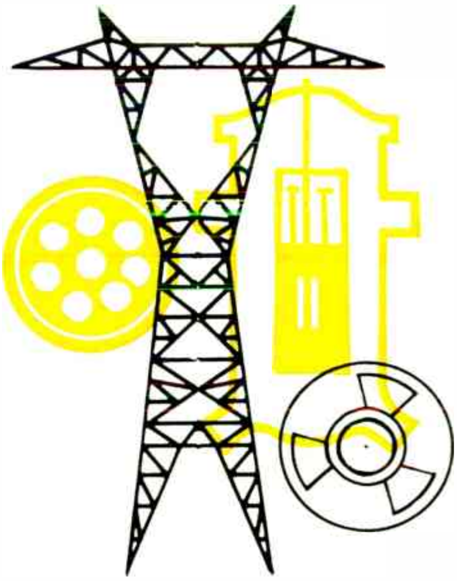
The importance of good cooling in superconductor application is now generally realized. Experimentation with various types of cooling systems and heat-transfer techniques will lead to a wide range of acceptable cooling techniques. Supercritical helium at high pressure is also being considered to increase acceptable heat-transfer rates by orders of magnitude. Gas cooling has already been shown to be a possibility. The pressure to develop large helium refrigeration systems below the λ point and in some NbZr magnets has already been reported. The elimination of resistive regions caused by flux jumps in short samples operated below the λ point has also been observed; these observations confirm the thermal nature of the problem.

Radiation. Little is known about the operation of different types of superconducting magnet in various types of radiation environment. More information is needed on the operating characteristics of the coils under irradiation and upon their anticipated lifetimes under integrated long-term exposure. The results will be colored by the specifics of design and the types of constructional materials used for insulation.

Radiation-induced defects appear to enhance the short-sample characteristics of some magnet superconductors. This enhancement continues with increasing defect density until a maximum short-sample characteristic is reached, beyond which further increases in defect density reduce the short-sample characteristic. Radiation-induced defects seem to anneal out near room temperature.

Radiation intensities necessary to induce defects usually cause sufficient nuclear heating within the superconductor to raise its temperature to or beyond the critical temperature. The problem of operation during irradiation, therefore, may reduce to limiting the superconductor exposure in the most favorable thermal environment to that which produces a minimal local temperature rise at any point in the winding.

The work described in this article was performed under the auspices of the U.S. Atomic Energy Commission.



A prediction of power system development, 1968 to 2030

By predicting the trend of future power system design some 60 years hence, we should be better equipped to solve some of the technical and sociological problems that the industry faces today

Alexander Kusko **Consulting Engineer**

Much planning for research and development of power system equipment to date has been on a short-term basis and has not taken into account the overall pattern for the industry. Long-range predictions, to the year 2030, are made in this article in order to help clarify the R&D picture and place the needs and times in proper perspective. A number of references are given, not to document in detail all significant predictions, but to show the trends of long-range thinking.

What will the power systems of the United States be like in the year 2030? How will power be generated and transmitted, and for what loads? In this article we shall make some predictions concerning the probable patterns of power system design beyond the extension of today's design practice. Such predictions, of course, are always of interest in themselves but, more important, they may enable us to formulate rational bases for the solutions of some of the technical and sociological problems that the power industry is facing today. However, rather than try to include the entire United States in our predictions, we shall only cover power system development for the heavily populated region of the Atlantic seaboard, typically, from Portland, Me., to Norfolk, Va., or for any other equivalent "megalopolis" in the country.

Assumptions

In making our power system design predictions for the period 1968–2030, we shall assume that:

1. The population will approximately triple. The population density will not change in the central areas of the cities, but will increase in the surrounding areas.
2. Bulk power generation will be primarily from nuclear-type power plants. New types of reactors and plants will be developed.
3. Annual peak load for each utility will approximately double each ten years.
4. Regulation of the electric power industry by the government at all levels—federal, state, and local—will increase.
5. Government interest in air and water pollution, nuclear fuel depletion, and preservation of scenic areas will increase.
6. The computer will become more dominant in power system operation and control.
7. Load management will become necessary to maximize the utilization of power system investment.
8. Higher power capability and less expensive cables will be developed to meet power system needs.

Bases for prediction

From a vantage point in the year 1968, we can make a rational prediction for the development of electric power systems well beyond the year 2000 by basing it on typical times required for the development, manufacture, and installation of equipment and the equipment's useful life after being placed in service.

Major equipment for a power system being planned during the late 1960s will not be manufactured and installed for at least five years and will require up to five years more for integration with the system. Based on past experience, we can say that power system equipment has a useful life of 30 to 50 years; equipment being planned and ordered today will still be in use in the years 2005 to 2025. In addition, equipment now under development, such as 1000-kV EHV equipment, new types of nuclear reactors, and high-voltage cables, will require an additional ten years for developing and testing, both in the laboratory and on a power system. This equipment will be installed and operating by 1985 and will be used in a power system up to the years 2015 to 2035.

Equipment now in the research stage for which an economic and technological breakthrough is not at hand, such as central-station magnetohydrodynamic (MHD) generators, superconducting cable, and solid-state circuit breakers, will entail at least ten years of further research, ten years for development and testing, and up to ten years for manufacture, installation, and integration into the system. This equipment will probably not be in service on a power system until the year 1995 and then should operate until the years 2025 to 2045. It is not likely that any major equipment will be in full operation on a power system by the year 2000 that is not presently under active research.

Present decade

The present decade, 1960 to 1970, is characterized by power system development through the construction of large generating plants and reliance on ac transmission networks at higher voltages to deliver power to the loads. Pooling arrangements and interconnections have been extended to achieve economy through diversity, to improve reliability, and to share capacity of new large plants.

Generation. The decade represents the peak of construction activity of fossil-fuel plants, particularly of mine-mouth plants. For site economy, reduction of air and water pollution, and fuel costs, the fossil-fuel plants are located outside of cities and serve their systems by EHV lines and networks. Toward the end of the decade, the falling costs of nuclear plants meet the rising costs of fossil-fuel plants and the latter become favored as the prime means for bulk power generation. Nuclear plants, built with boiling-water and pressurized-water reactors, approach 1000-MW module size and 2000-MW plant size by the end of the decade. Because of radiological hazards, they are built away from cities, but close to cooling water, and require generally shorter lines to deliver power to their loads.

Hydroelectric plants continue to be built for flood control, water supply, irrigation, and power generation. Pumped storage plants are proposed and built to provide off-peak energy storage facilities. Development of Canadian hydroelectric power at Hamilton Falls and other remote locations is planned.

Generator packages have been developed using, at first, diesel, then gas-turbine, prime movers with ratings up to 160 MW. These self-contained units are installed near existing generating plants to provide peaking service and plant start-up power, and to improve reliability. They are also developed for isolated operation in buildings where they provide heating and air conditioning as well.

Transmission and distribution. Construction of short and long compensated transmission lines proceeds rapidly for both internal system transmission and interconnections. Voltages are increased to 500 and 765 kV ac, and to 1000 kV for dc equipment, to increase power-carrying capability.

Distribution construction proceeds in the direction of higher voltages up to 69 kV, and also toward underground facilities. Much new installation in urban areas is made underground, particularly where the costs are shared. Existing systems are converted to underground in areas of urban redevelopment, where streets are widened, or where necessitated by local ordinances.

In New York City ac cable has been developed to 345 kV and installed. Development work is proceeding on 500-kV ac pipe-type cable, direct-burial cable to 230 kV, and sodium-filled cable. Cable for dc service is being developed to 500 kV. Research is being carried out on new dielectric materials, methods for cost reduction, cooling methods, and new methods of transmission in a pipe.

System design. Power systems are operated with computer assistance, but with decision making by man. Analog and digital computers are employed to regulate tie-line flow and to obtain most economical system dispatching. Individual power plants are highly automated and employ computers for starting and operating.

System protection is based upon air-blast and oil circuit breakers and electromechanical relays, but solid-state relays for special functions are being introduced. Initial work is being done on system-wide coordinated protection; however, the bulk of the protection is confined to lines, buses, or specific equipment.

Massive blackouts in 1965 and 1967 have forced system designers to evaluate interconnections, spinning reserve allocation, system operation, and standby generation, and have resulted in the adoption of load-dropping techniques by underfrequency relays and in more emphasis on strong interconnections.

Stringent regulations are passed against air and water pollution, particularly in cities. Ordinances by local, state, and federal government on rights of way force underground construction in cities, suburbs, and scenic areas. Utility corridors are being considered in urban areas for all overhead and underground services.

Loads. The loads are typically lighting, heat, and air conditioning for homes, offices, plants, and stores; street lighting; transportation; manufacturing equipment using motors and furnaces; and chemical and metal processes.

Prediction for decade 1970 to 1980

The years 1970–80 see a continued growth of power systems following the same design concept as at the end of the previous decade. This decade is characterized by further interconnections at higher voltage, larger nuclear generating plants, long-distance dc systems, and higher-voltage distribution systems having a larger percentage of their facilities underground.

Generation. Practically no plants using coal are built. Nuclear plant construction accelerates, using boiling-water and pressurized-water reactors for modules up to 1500 MW and plants up to 5000 MW. Nuclear fuel costs spur development work on breeder reactors, heavy-water reactors, and the fusion process. Nuclear plants are still built away from populated areas, and frequently require cooling towers to reduce water pollution.

Generator packages using gas turbines in modules to 500 MW become highly developed in regard to efficiency, reliability, and availability. Units are installed for peaking, reserve, and isolated generation, and to supply additional services, such as heating and air conditioning.

There is a surge of portable power equipment usage, centered on the battery-powered electric automobile. The battery-and-motor combination, with its reduced smoke and noise, takes over many of the one- to ten-horsepower functions of the small gasoline engines. Development of central-station MHD generators continues on a laboratory scale.

Transmission and distribution. The pace of overhead line construction reaches its peak in this decade. Working voltage levels rise to 1000 kV for ac lines and ±1000 kV for dc lines. Extra-high-voltage networks are required for transmitting power from large nuclear plants and for interconnections. Existing rights of way become crowded and new rights of way expensive, forcing utilities to utilize multicircuit towers, corridors, and cable in and around cities.

The power requirements on distribution circuits force voltages upward to 138 kV, and the legal restrictions on rights of way in populated areas force the circuits underground. By 1980, nearly all new distribution circuits in cities are placed underground, and old circuits are transferred underground as the interiors of cities are rebuilt. System design trends toward only primary voltage distribution to transformers located at the individual loads.

Cable development work results in 500-kV ac cable and 1000-kV dc cable for termination of overhead lines and as independent circuits. Direct-burial cable for distribution up to 138 kV has been developed and is in use. Intense work is carried out on compressed-gas insulated cable, superconductor cable, and artificially cooled cables to provide economical high-power links.

System design. Computers are used to control and optimize the operation of individual nuclear plants, and toward direct computer control of all plants, interconnection flows, and some loads for least-cost operation. Computers are programmed for decision making during emergency conditions, and to monitor incipient conditions that could lead to massive system failures.

Air-blast, oil, and some vacuum breakers are used for system protection. Solid-state breakers are being developed for special applications. A dc breaker is being developed. Electromechanical protective relays are being displaced by computer-type equipment, which looks at the complete system rather than at subsystem portions. Protective systems are being designed not only for fault isolation, but to assure transient and dynamic stability of the system.

The Federal Power Commission assumes regulation of interstate lines and sets requirements for power-line undergrounding in populated and scenic areas. State and local ordinances require that all overhead telephone and electric lines within populated areas be underground by the year 2000. Stringent air, water, and electromagnetic pollution requirements prevail.

Loads. Load growth occurs in all of the categories of the previous decade. In addition, off-peak charging loads for electric automobiles and other portable equipment rise. Space-travel and airport complexes, highway systems, and high-speed railroads require large amounts of power. Electrochemical and electrometallurgical loads increase for material manufacturing and processing.

Prediction for decade 1980 to 1990

The decade 1980–90 sees a change in the system design pattern of the 1960s and 1970s. By 1987, system peak loads have risen by a factor of four from 1967, and will rise to eight times by the year 1997. These load projections, combined with growth of population density, make it apparent early in the decade that system design must be altered to service the predominantly urban area.

System design follows one or more of the following routes to meet demand:

1. Nuclear reactor plants for operation within a city and with self-contained cooling facilities. The plants are of either high-rise or underground construction. New loads are served by direct high-voltage ac cables from plant to load centers.
2. Nuclear plants located up to about 95 km (60 miles) outside densely populated areas and using natural or artificial cooling facilities. Power is transmitted to centers of populated areas by means of high-voltage, high-capacity cable, such as compressed-gas insulated-type cables, placed in underground ways.
3. Nuclear plants located where convenient and using dc cables and terminal facilities for transmitting power underground into centers of populated areas.

The choice of development route depends upon economics, technological breakthrough, and radiological hazard control. The conventional ac cable with its requirement for compensation every 30 km and limitations on ratings is probably not a realistic means for transmitting power from plants outside the population centers.

Generation. Nuclear plant construction continues with the first installations of heavy-water reactors and experimental installations of breeder reactor plants. The module size for the turbine and generator unit reaches 2000 MW; complete plants reach 10 000 MW. Pool sharing of module capacity is required less as each system is able to absorb the full output of each new module to match its load growth. Reliability and reserve are obtained by module redundancy within a power system. Work on nuclear-fusion production of power is intensive. Planned industrial–residential cities are built around nuclear plants where the reactors supply a basic need for each industry, such as radiation or heat for food processing, chemical reactions, and the manufacture of new materials.

Generator packages using gas turbines or equivalent sources are installed in load areas for peaking and supplemental generation and are operated under computer control. Installations of packages for power, heat, and air conditioning are numerous in buildings and load centers.

Rechargeable or directly fueled power sources driving electric motors assume functions previously fulfilled by engines up to 100 horsepower. A plant is built using a direct fuel-to-electric-power conversion (MHD). Exotic energy sources, using heat, solar radiation, and nuclear materials, furnish isolated power for communications, data transmission, space and lunar stations, and alarms. Development work continues on large and small power sources.

Transmission and distribution. The pace of overhead transmission-line construction is reduced toward the end of the decade; lines are restricted to existing rights of way and utility corridors and new ones are built only for long circuits to remote plants and interconnections. Working voltages have leveled off at 1000 kV ac and ±1000 kV dc for overhead lines because of clearance requirements for air insulation. Existing overhead lines into and around cities are being converted to cable circuits. New, centrally located plants are serving load centers by direct ac cable circuits.

Distribution systems utilize voltages up to 230 kV in cable and serve load-mounted transformers directly. Limited distribution is provided at the lower utilization voltages. In densely populated areas, all new construction is underground and much of the existing overhead distribution circuits have been converted to underground cable.

Development of 1000-kV ac cable is near completion and test installations have been made on superconductor and gas-insulated cables. Interconnections are being made with ±1000-kV dc cable to increase capacity, cover the distances, reduce fault currents, and improve stability.

Direct-burial 230-kV cable has been developed for distribution.

System design. Computer-type equipment controls all dispatching and protective system functions during normal and abnormal conditions, removing man from the decision-making loop. Certain process loads are being developed for operation under the control of the utility to insure the most economical loading of the system.

Solid-state switches are involved in test installations but reliance is still placed on air, oil, and vacuum circuit breakers. Cable is installed integral with terminal and protective equipment. Fault detection and switching are provided by a central computer, which also monitors stability. Initial work is carried out on cable-monitoring equipment to detect incipient breakdown and carry out protective measures.

Development work and testing continue on dc converter terminal equipment suitable for dc cable systems between nuclear plants and load centers in densely populated areas. Preliminary design is made of nuclear plants operating with higher turbine-generator speeds and frequency to feed into dc systems.

In moderately populated areas of 500 persons or so per square kilometer, conversion of all overhead circuits to underground is required by ordinance within 20 years. Fuel burning of any kind—industrial, domestic, or for transportation—comes under strict regulation. Stringent federal control is established for power system practice, including reliability and development work.

Loads. Heavy load growth is experienced for light, heat, and air conditioning of large domed areas for shopping, industry, and recreation; for space travel and satellite terminals; and for electrical processes associated with agriculture, chemical, metallurgical, and other industries. Cities spread and merge into large homogeneous population areas; utilities reorganize boundaries to accommodate more logical service of load areas.

Prediction for decade 1990 to 2000

Because of the tremendous growth of load, the decade 1990–2000 is marked by the transition of sprawling power systems into compact load-area subsystems having balanced load and self-sufficient generation. Interconnections have less function but are maintained for existing plants that depend upon them. Plants built up to the year 1960 are retired. Extensive computer control is employed for all system functions. All new circuit construction, except in special cases, is in underground cable.

Generation. No new fossil-fuel plants are built for base load. Existing plants built between 1960 to 1990 are operated as required. Gas and coal products are used for gas-turbine-generator packages or fuel cells.

Nuclear plants are being built using heavy-water reactors and are located either in population centers or adjacent to them. A design is evolved for a high-rise plant with integral cooling towers or closed-cycle cooling for locations where water is unavailable. First commercial installation is made of breeder reactors. Power is distributed directly to major loads and distribution centers by cable. Nuclear plants supply additional functions, such as water desalinization, product and food irradiation, and heat. Plants are under local and central computer control and are practically unmanned. Industrial–residential cities continue to be built up around nuclear plants.

No new hydro plants are constructed; existing plants are operated. Plants built prior to 1960 are retired. Pumped storage plants requiring overhead transmission lines are being retired in favor of gas-turbine-generator packages at the load centers.

Self-contained energy-source packages using other than gas turbines are being developed; fuel cells and energy storage elements are used. Extensive use of these units for peaking, reserve, and local generation and for the supply of other services is well developed. Their operation is either synchronized with that of the system or is isolated under central computer control.

A bulk-power MHD plant is tested on a power system. Large fuel cells and other direct electric conversion equipment are put into operation. New small power sources for ground and space purposes are developed. All movable and moving equipment uses portable rechargeable sources or fuel-using sources of electric power.

Transmission and distribution. Few new overhead lines are being built. Lines built to the year 1970 are being dismantled or consolidated at a higher voltage level but those built after 1970 are still in service; they include most of the 500-kV and all of the 765-kV and 1000-kV lines. These remaining lines are well away from densely populated areas but are brought into these areas by cable. New circuits are being built as ac cable up to 1000 kV within a system and as dc cable circuits for interconnections and for bulk transmission from plants to load areas.

Nearly all overhead distribution circuits are converted to underground by the year 2000; direct-burial cable is now cheaper than overhead distribution lines. Power and voltage levels are too high for overhead circuits in cities. Various classes of loads receive appropriate service. Generator packages and load-control systems keep major high-voltage cable links and base-load plants fully loaded.

System design. Load areas are sectionalized, fed from cables, and under computer control of loading, protection, and stability. Bulk power generators, controllable loads, and local generator packages are operated to optimize economy.

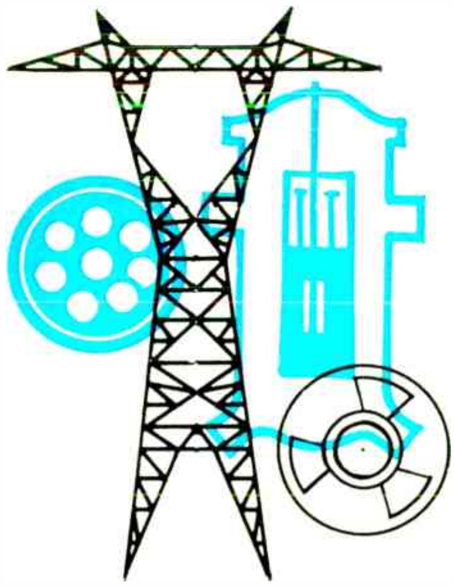
Protective systems are designed for high-speed interruption of faults using solid-state switches. Monitoring systems detect incipient fault conditions in cables and equipment by measuring ionization currents and other techniques. All switching on the system is under control of a computer. Continuous power to loads is assured by circuit and source redundancy and fast cable and equipment switching for fault clearing and rerouting.

Loads. Large portions of cities are domed and use electric power for maintaining climate. Space and earth travel utilizes electric power for launching vehicles, terminals, and communications. Industrial processes for the manufacture of food, materials, goods, and products require electric power. Nearly all mobile and portable equipment uses electric power.

Prediction for 2000 to 2030

The period 2000 to 2030 must be considered first in terms of the predicted electric power load. Assuming that the load doubles every ten years, in the year 2000 the system load will be ten times its 1967 value, and in 2030 it will be 80 times its 1967 value.

The load densities and power levels make it apparent



that unless superconductor cables are developed generation and load must be physically close to reduce the need for transmission. The pre-2000 solution of using increasingly high-voltage transmission lines to supply such loads will not work. Means must be found during this period to build large nuclear plants directly in densely populated areas to provide electric power and other services.

Generation. Nuclear plants using heavy-water and breeder reactors supply all of the bulk power directly by cable to load centers. Plants designed for installation adjacent to or in population centers use integral cooling towers or closed systems where necessary, and provide other functions, such as radiation, desalinization, heating, and cooling. The first fusion plants may be in operation in the 10 000- to 25 000-MW range along the seacoast to provide electric power and fresh water.

In the decade 2000 to 2010, all light-water plants and interconnections built during the period 1960 to 1970 are retired; during the next decade all plants and lines dating back to 1970–80 are retired; by the year 2030, all light-water nuclear plants and the massive interconnected 500-kV and 765-kV lines and grids that were built to the year 1990 have been retired and dismantled. The nuclear reactor plants of the 2000–2030 period operate entirely under computer control and at optimum power level. Loads and peaking generator sources are controlled to keep the reactors and their power cables under maximum steady-load conditions. Large hydro plants and their lines, built by 1990, operate as dictated by economy in the 2000-to-2030 period, but are finally retired by the year 2030.

Generator packages using gas turbines or other direct converters are operated at load centers to supply variations of power from the base level. They also provide such services as reserve, heat, and air conditioning, under control of a central computer, and they utilize fossil or fission fuels or recharge from the bulk power generators off-peak. They are capable of operating either isolated or synchronized to the system.

Magnetohydrodynamic generators are used for bulk power with processed fossil fuels. Fuel cells and batteries are widely utilized for portable and mobile equipment. Small nuclear-reactor generator packages are used for space stations and for lunar and planetary power. Fuel management becomes the criterion for selection of electric power generation methods.

Transmission and distribution. Practically no overhead lines are built after 2000. Power is carried from new generating plants to load centers by high-voltage cable; links between plants are in cable and are heavy enough to carry the power of at least one module during emergency and maintenance periods. Interconnections with other systems are by dc cable.

Practically all distribution systems are underground by the year 2000. Loads are graded for required reliability and are supplied with redundant cables as necessary. Energy-source packages at the load buses provide for the variations in load from the base-load value and also provide reserve generation in case of cable failure. Certain loads are managed to insure nearly constant levels of system power loading. There is no low-voltage distribution, only feeders to transformer banks at load sites, or direct utilization of the feeder voltage. Circuit interruption and monitoring systems switch loads from faulty cables to other sources without power interruption.

Cable development after the year 2000 is directed at two classes of cable—one for bulk power transmission from plants to load centers and interconnections and the other for distribution purposes. The bulk power cable may be required to carry 5000 to 10 000 MW per run and utilizes superconductors, compressed-gas insulation, artificial cooling, or some other technique to achieve very high power-flow densities. Fast switching for fault isolation and power rerouting, as well as means for fault monitoring, make cable installations highly reliable. Bulk power cables, both ac and dc, approach the 2400-kV level. Distribution cables approach 500 kV. Cableways are placed in populated areas for ease of cable placement, service, and modification.

System design. First, the systems are designed to operate in self-contained load and generation subsystems, which are loosely interconnected. Second, all controls and supervision of the system are handled by computer, which also maintains the system in optimum condition. Third, loading as well as generation is managed by the utility to utilize capital equipment to its fullest. Fourth, practically all new transmission and distribution loads are carried by cable.

Federal and local regulation and standardization of equipment and operating means continue after the year 2000. The electric power industry as we know it in 1968 may well disappear in favor of a vertical industry combining generation and utilization.

Loads. Load growth continues for environmental control, materials processing, manufacturing, travel, and communications. Nuclear reactors and other fuel sources are utilized directly without conversion to electric energy as an intermediate step. Further integration and control of the generation and utilization equipment are carried out to achieve the most economical means of reaching the end requirement.

Conclusions

A study of the future of the electric power system in a populated area such as that along the Atlantic seaboard indicates the following developments by the year 2030:

1. The area will develop in terms of an EHV trans-

mission grid reaching 1000 kV, with nuclear plants located outside cities until about 1990. It will then develop further in terms of large in-city nuclear plants supplying subsystems by means of high-voltage underground cable, where the subsystems are loosely linked.

2. The most important generation requirement by the year 2000 will be for a plant that can be located in a populated area, operate independently of cooling water, and generate 10 000 to 25 000 MW.

3. Underground systems, such as compressed-gas insulated cable, will be required by the year 2000 for bulk transmission of 5000-MW blocks over distances of 95 km. Distribution cable up to 230 kV will be required to service populated areas.

4. Direct-current cable and converter systems will be used beyond the year 2000 for supplying densely populated areas from peripheral plants and for system interconnections.

5. All distribution systems will probably be underground by the year 2000 and all transmission lines will disappear by the year 2030. The lines will no longer be required for long system-to-system interconnections, nor for carrying power from remote plants to load centers.

6. Power systems will be placed under progressive computer control for dispatching, fault detection, protection stability, and reliability of customer supply. The limiting factors on progress will be the slow development of sensors and techniques.

7. Operation of systems will change from a peak-and-valley load cycle to steady loading of major equipment, as companies develop supplementary peaking energy sources, and as they assume control of the power demands of manufacturing processes, such as water desalinization.

Realistic planning for power system expansion using techniques beyond the concept of EHV grids will guide new equipment development, and will serve to place on a rational basis the selection of rights of way and generating plant sites, and the formulation of schedules for placing underground all of the present overhead facilities.

Concerning the references

Anyone who studies the problems of providing services to a heavily populated region such as that along the North Atlantic coast should refer to Gottman's original work on "the megalopolis,"[1] and subsequent references. A useful recent reference is the special issue of the *Technology Review* on "Cities in Crisis."[2] A specific reference to services for a planned city is given by Gipe.[3]

Typical references showing extrapolations into the future of past and present growth rate, equipment and system design, and methods for serving load areas, are given by representatives of the government, electric equipment manufacturers, and the utility industry. The "National Power Survey"[4] shows extensive projections of loads, miles, and types of lines, and methods of generation to the year 1980. Alexander, Johnson, and McConnell[5] give projections to the year 1970 of the ratings of EHV transmission equipment. The report on "Power Supply for New England, 1973–1990"[6] shows an extrapolation of present design methods to serve a load area to the year 1990.

Sample references that reflect equipment needs and designs for the future are those by Starr[7] on nuclear reactors, by Garwin and Matisoo[8] on superconducting cable, by Philp[9] and Fukuda[10] on compressed-gas insulation, by Cohn[11] and Friedlander[12] on computers, by Tsu[13] and Brogan[14] on MHD generators, and by Friedlander[15-17] on esthetics, pollution, and loads; see also the *Conference Record* of the 1967 Power Industry Computer Applications Conference.[18]

Equally as important as references to technical factors in determining the future course of power system development are references to legislative and regulatory rulings. Decision no. 73078[19] of the California Public Utilities Commission, ordering the utilities of the state to assign specific funds to conversion of overhead distribution lines to underground, may become the pattern for other states. Proposed legislation on underground conversion that is typical for several states is Massachusetts bill H984.[20] Of the various pieces of proposed federal legislation on electric power systems, the "Electric Reliability Act of 1968"[21] is receiving major attention.

The cited references show that long-range prediction of power system development must include not only technical developments, but the sociological changes and the impact of federal and state regulations as well.

REFERENCES

1. Gottman, J., *Megalopolis, The Urbanized Northeastern Seaboard of the United States*. Cambridge, Mass.: The M.I.T. Press, 1961.

2. "Cities in crisis," *Technol. Rev.*, vol. 70, Jan. 1968.

3. Gipe, A. B., "Planning a new city—Columbia," *IEEE Trans. Industry and General Applications*, vol. IGA-2, pp. 423–430, Sept./Oct. 1966.

4. "National power survey, pt. I," Federal Power Commission, Washington, D.C., 1964.

5. Alexander, G. W., Johnson, J. B., and McConnell, A. J., "Trends in extra high voltage transmission equipment," *Proc. Am. Power Conf.*, pp. 988–997, 1965.

6. "Power supply for New England, 1973–1990," Ebasco Services, Inc., Feb. 1967.

7. Starr, C., "Chauncey Starr looks at the U.S. national reactor program," *Nucleonics*, pp. 46–51, June 1967.

8. Garwin, R. L., and Matisoo, J., "Superconducting lines for the transmission of large amounts of electrical power over great distances," *Proc. IEEE*, vol. 55, pp. 538–548, Apr. 1967.

9. Philp, S. F., "Compressed gas insulation in the million-volt range; a comparison of SF_6 with N_2 and CO_2" *IEEE Trans. Power Apparatus and Systems*, vol. 82, pp. 356–359, June 1963.

10. Fukuda, S., "EHV cables with compressed SF_6 gas insulation," *IEEE Trans. Power Apparatus and Systems*, vol. PAS-86, pp. 60–66, Jan. 1967.

11. Cohn, N., "The automatic control of electric power in the United States," *IEEE Spectrum*, vol. 2, pp. 67–77, Nov. 1965.

12. Friedlander, G. D., "Computer-controlled power systems," *IEEE Spectrum*, vol. 2, pt. I, pp. 61–81, Apr. 1965; pt. II, pp. 72–91, May 1965.

13. Tsu, T. C., "MHD power generators in central stations," *IEEE Spectrum*, vol. 4, pp. 59–65, June 1967.

14. Brogan, T. R., "MHD power generation," *IEEE Spectrum*, vol. 1, pp. 58–65, Feb. 1964.

15. Friedlander, G. D., "Esthetics and electric energy," *IEEE Spectrum*, vol. 2, pp. 46–55, June 1965.

16. Friedlander, G. D., "Airborne asphyxia—an international problem," *IEEE Spectrum*, vol. 2, pp. 56–69, Oct. 1965.

17. Friedlander, G. D., "Science and the salty sea," *IEEE Spectrum*, vol. 2, pp. 53–64, Aug. 1965.

18. *Conference Record*, IEEE Power Industry Computer Applications Conference, Pittsburgh, Pa., May 15–17, 1967.

19. "Investigation on the Commission's own motion into the Tariff Schedules, Rates, Rules, Charges, Operations, Practices, Contracts, Services, and Aesthetics and Economics of Facilities of All Electric and Communication Public Utilities in the State of California," Decision no. 73078, Public Utilities Commission of the State of California, Sept. 19, 1967.

20. "An Act Relative to Poles and Overhead Wires and Associated Overhead Structures," H984, The Commonwealth of Massachusetts, 1968.

21. "Electric Reliability Act of 1968," proposed.

New faces of Eve: women in electrical engineering

With the increasing shortage of engineers, it is time that we utilize the capabilities and creative talents of our women, and recognize that they constitute the missing half of our technical potential

Irene C. Peden *University of Washington*

The gap between the number of highly trained technical people and the positions to be filled is widening year by year. There has been much soul searching on the parts of the profession and of education as to where the fault lies, but one important reservoir of talent remains largely untapped—the woman engineer. This article examines the motivations of the woman who selects engineering as a career, and discusses her problems and potential in a field that has always been associated with the masculine image.

There is plenty of room at the top in electrical engineering. Furthermore, it would seem that we have reached saturation as to our traditional pool of talent, and that the number of promising young men entering the field is not going to increase much more. How can we hope to fill the gap between the number of people with a high degree of technical training and the number of challenging positions awaiting them? The vitality and continually expanding horizons of our field have kept it in the mainstream of technical thought for many years. If it is to prosper equally in the future, we need to believe as a profession that women are the missing half of our technical potential, and that they really have the talent for creative contribution. We should be asking what the woman engineer is like and what motivates her. How does she view her career choice? What is her professional future? How can we encourage a significant increase in the number of women among us?

Woman engineers are a minority group in the field, to a greater degree than can be justified on objective grounds. The results of basic aptitude tests given to mixed groups of high school students have indicated an aptitude for engineering in 6.3 percent of the boys and 4.2 percent of the girls.[1] Thus if all the girls whose natural talents ought to lead them into engineering were actually following through, we would have two women for every three men in the field. Instead, less than one percent of all the engineers in the United States are women.[2] In 1966 more girls were enrolled in chemical engineering than in any other engineering discipline (electrical placed second, with civil and mechanical following in that order).[3,4] In terms of the realities of the working situation, one would expect electrical engineering to be their first choice; instead, they are attracted to an area that does not appear to hold the most promise for them. Perhaps the trend need not be a long-range one if we recognize the existence of a problem, and begin to work at communicating with students and with the adults who influence their career decisions. In order to do this effectively, we must be convinced that women are needed and wanted in our field, and that any recruiting we do is realistic.

What is she like?

Any woman engineer who has made it beyond the second-year college level is certainly serious about her choice, and willing to work hard to learn and to apply the material. She enjoys the challenge of engineering problems and the satisfactions of solving them. Some women engineers may be quite competitive, whereas others seem pretty low key about relative performance, attitudes that are equally present in the male engineering population.

A coed's ability to do the classroom work is no different from that of her fellow students, given that she has made her curriculum decision on the basis of aptitude and interest; it follows that she can do the theoretical work of her field as well as anyone else. The girl who needs assistance with lifting or moving heavy objects is certainly not limited in the choice of electrical engineering jobs; technicians are usually available for these chores, which are rarely performed by the professional engineer, even when he is physically able.[5]

When there is precision laboratory work to be done, many women have the edge; they generally outperform men where fine finger dexterity is concerned. They may be a little slower to catch on when using tools or working with mechanical equipment, should that be required occasionally, but if so, a lack of basic ability may not be the explanation. Little girls take the alarm clock apart less often than their brothers—not because they wouldn't like to know what makes it tick, but because they may be sensitive to what parents and playmates think is appropriate activity for little girls. It follows that they are less sure of themselves later on in this respect, and less likely to take the initiative in a group. Some women are limited in mechanical aptitude, of course, but even that presents no particular handicap to the electrical engineer, who relies on mathematical techniques and physical reasoning. Today's successful engineer is good at mathematics, with a flair for intellectual pursuits.

Without question, the most famous woman engineer living today is Dr. Lillian Gilbreth, the mother of 12 in the well-known book *Cheaper by the Dozen*. This pioneer industrial engineer has made use of her training in motion study and scientific work methods to help the handicapped and disabled.[2,6] And if one thinks about it, it seems to be no accident that a woman was the first to see this warm and human application.

Speaking generally, women tend to be more people-oriented than men, as evidenced by the professions that have grown up around their special skills—such as nursing, teaching, and social work. They have a special sensitivity to the needs of others, and often outperform men where patience and efforts toward harmony are concerned. They are often more effective in communicating with other people, too, and can bring new dimensions to the "masculine" professions, as Dr. Gilbreth has done.

Psychological studies suggest that as the minority group in the engineering world, women may have some unique contributions to make, due in part to the fact that they have a somewhat different way of looking at the problems.[2,7] This ability to see things differently, to reweave the fabric of one's discipline, is the basic ingredient of creativity. Perhaps we should view women's potential in any of the traditionally masculine professions in these terms, over and above the evidence that they can meet existing technical challenges with the same effectiveness as their male colleagues.

Then why is there this great discrepancy between the numbers of women who can qualify on the basis of inherent ability and the numbers actually found in engineering? Much of it is due to the discouraging influence of parents, teachers, counselors, and friends during the important early years when career decisions are made. It often happens that girls find out too late that engineering is one of the most promising of their choices, when the time has passed when they could still organize their preparatory courses to meet the entrance requirements of the engineering colleges. It is unfortunate that the image of the engineer (hard hats, construction, and heavy machinery) lags so far behind the facts in the public eye, and does not apply to electrical engineering at all. When the public begins to realize that electrical engineers are concerned with solid-state electronics, computer design, communication theory, radio astronomy, and bioengineering, to name a few of our areas, surely the old-fashioned masculine image will blur and fade. Perhaps then we can look forward to more encouragement of talented girls by the well-intentioned adults who influence their career decisions.

The woman who identifies herself with one of the professions historically populated by men (engineering, medicine, law, architecture, etc.) is viewed as something of a maverick in this culture. However, a psychologist has this comment: "Should a woman display interest in any activity traditionally considered masculine, she is sure to be accused of rejecting her 'female role.' Our data suggest that the contrary of this assumption is likely to be true. The girls that showed relatively high enthusiasm for the activities of the traditional female role also tended to show more liking for some traditionally male activities. This suggests that receptive and positive people are likely to be receptive and positive about many kinds of things, without regard to the artificial boundaries of traditional sex assignments."[6,8]

Although the maverick is by no means a masculine woman, she has certainly selected a rigorous and demanding direction in which to turn much of her attention. She has at some time in her life prepared herself with a very specialized training curriculum—perhaps a long one if she has done graduate work. It appears significant that the "feminine" professions generally require less formal training for the top positions than the "masculine" ones, a natural adaptation no doubt to the different life patterns of men and women. It is often true as well, and probably for the same reason, that the successful person in a man's field puts in longer hours to ensure that success than does his or her counterpart in an area where the majority of colleagues are women. There is no implication here that women are less capable of hard work than men. Many working women have family responsibilities, however, and their time is thereby divided to a greater extent.

There must be nearly as many reasons for their career choice as there are women in the maverick category, but surely the common denominator is a mixture of interest, ability, and enthusiasm, just as it is for men. No other career motivation is realistic; no other would sustain the candidate beyond the freshman year of training.

The well-motivated girl who combines these qualities will be successful in college, which serves as a valuable testing ground in many ways, including the technical. By the time she graduates, she will know what it is like to be a lady engineer, although she will still have as much to learn about the practice of engineering as any other student. She has encountered a certain amount of resistance to her presence in the classroom and laboratory, and she will find some in the industrial world, in an occasional supervisor, a few colleagues, and in some service personnel from time to time. But she has also experienced the fact that for every situation that was difficult, there were several others in which people went out of their way to be helpful and encouraging—and this will be equally true on the job. She and her prospective employers can be confident that her career holds further surprises of this sort for her. She has had the opportunity in college to know where she stands academically relative to her classmates, and should be able to make a realistic judgment about the type of work she should choose in order to be productive and happy.

She has already come to terms with her high degree of visibility. No girl in engineering college can ever suffer from the multiversity student complaint of feeling like an IBM card; no practicing woman engineer has any illusions about her inability to fade into the background. Instead, she feels spotlighted much of the time, and this is not entirely comfortable. Her position can be rather lonely at times, even for the self-sufficient. One girl who obtained a B.S. degree in electrical engineering from the University of Washington several years ago made the remark that the boys were pretty standoffish until they were sure she was serious about her studies; then she became an unusual member of the group.[2,5,6] This pattern is an effective preview of the professional challenges ahead. Getting the job in the first place is usually the major hurdle for the woman engineer. Once established, she becomes a valuable, though still unusual, member of the group. Her salary level is the same as that of a man with equal capability, and so is her workload.

Women with administrative ambitions are well advised to be conformists in their career choices, and to stick to such classical fields as teaching, nursing, and home economics. Because these are populated largely by women, and the turnover rate is fairly high among the younger ones while their marriages are new and their children young, the dedicated woman professional who is good at her work and sticks with it has a good chance to rise to the top. The maverick, on the other hand, had better like the work for its own sake, for her chances to advance to top administrative positions are not bright, even in 1968. Women engineering executives do exist,[2] but it is worth pointing out that although many rise to supervisory positions, more do not. The situation can only improve, as more and more talented women enter and stay in the engineering profession to demonstrate their capability. The woman who can handle more responsibility and who wants it will identify herself to supervisors and/or recruiting personnel.

Women engineers do well to select new fields where there are few preconceived notions about what is masculine and what is not. Engineering presents a wealth of fascinating choices in all categories, and there are more exciting possibilities for today's engineering student than at any time in history.[9] Many of our own specialities are new; electrical engineering covers a broad spectrum of technical endeavors, and promises to encompass more as it continues to expand its horizons.[10] With more opportunities to offer, we should have a bigger welcome mat out than any other profession to the new type of woman emerging in our culture. We are in a position to encourage her to help populate our new career fields as she has already done elsewhere. Psychiatry, one of medicine's newer fields, is not considered as masculine as surgery, and many women physicians excel in this specialty. Consider the numbers of women who have poured into computer programming in recent years, now that there is interesting work to do with a bachelor's degree in mathematics. They belong in computer science, too, at a higher professional level. Why should computer science be viewed as masculine? Or quantum electronics?

Although the girl graduate in electrical engineering will have her share of job offers at the same salary level as men of equivalent scholastic standing, she may find employers reluctant to hire her on the grounds that she may quit as soon as they have trained her because their file clerks often do. If she has enough insight, she can recognize the weakness of this argument. The same strong interest and motivation that carried her through the engineering curriculum will give her a high degree of job stability. Enlightened management people know as she does that her reasons for working are not the same as those of file clerks. Of course she may change jobs at some future time—any new employee may, and for a variety of reasons. There is always a calculated risk in hiring new personnel.

Her life as a professional

Women are in the working world to stay. Given that, what is important in their environment? For most, the intrinsic satisfactions of the job are more highly motivating than sheer toe-to-toe competition. The majority of women really do not want to be rivals of the men in their fields. They just want to do the jobs for which they are trained. They, like all other engineers, should be encouraged to do those things they can do best. Some are more likely to succeed where interest in people is an important ingredient of the job; others are well suited to research and various kinds of analytical and laboratory work.

What can her profession expect of the woman in a man's field? Obviously, the answer is technical competence and her best effort. In addition, there are some subtle considerations that might profit from a good airing. The aware maverick must learn that no criticism of her as a person is implied when her ideas are subjected to critical analysis, as they inevitably will be when she is contributing directly to a professional effort. A magazine cartoon comes to mind, by way of illustration. A couple is shown at the dinner table with a steaming casserole dish between them. He asks what is in the stew, and she replies, "Why? What's wrong with it?" It is often true that women must give special attention to overcoming this tendency. It is a pretty secure professional woman who can share her ideas openly as they develop. It is a pretty secure male colleague or supervisor who can be helpful and reassuring in this regard.

It is important that she be as objective as possible with colleagues or clients, although such positive qualities as good humor and understanding are always in order. It is a rare "pro" who displays inappropriate feelings on the job, contrary mythologies and mystiques notwithstanding. The novice can be advised to acknowledge that her culture permits it—and then pass up the opportunity. She may be happier in her apprentice years, too, if she consciously reminds herself from time to time that men, by and large, do not give or expect praise on the job to the extent that women do. Her colleagues and supervisors might be favorably impressed with the results of communicating their approval directly to her when she is doing her work well.

The successful professional person is necessarily assertive in the sense that the best interests of one's project, students, technicians, etc., must be looked after. At the same time, the woman who appears overtly aggressive is usually headed for trouble with colleagues and employees. Balancing these important factors in a man's world can result in some personal wear and tear, although it represents a good challenge to the mature person. Any girl who has successfully completed four years or more of engineering college knows about it. She may not handle it as smoothly at first as she will later on when she has had more life experience, but there should be no doubt that she has already learned much about the need to balance professional competence with feminine tact and charm to produce a valuable set of job qualifications. There should be no doubt about her inherent ability or willingness to develop expertise in this regard.

The first women to enter medicine, law, engineering, etc., must have been pretty stubborn and aggressive—"unfeminine," in other words. The pioneering is done, however, and it is not only possible, but an accomplished fact, that the mavericks in today's fluid society can present a softer image without any sacrifice in professional standing. As the number of such professional women increases, so does the number of role models on which high school girls, parents, and teachers can base their evaluation of appropriate careers for women. Surely the time is coming when we will no longer sort career women into "maverick" and "conformist" categories. Already the traditional boundaries of male and female roles in society are breaking down. By their very modes of dress our teenagers may be telling us that their generation accepts equality of the sexes; they must be expressing something.

The woman engineer is just as likely to marry and have a family as anyone else.[11] She is more likely than the average woman to go back to work when her children are older, because of that same strong interest and drive.

Exceptional job stability in her age bracket is a matter of record. She will have to try to keep up during her years at home, which can be quite difficult. Alternatively, she will need to retrain herself before re-entry, possibly by returning to school for a graduate degree or by means of continuing education courses. She can, at the convenience of herself and her family, then offer the valuable combination of mature judgment and up-to-date technical knowledge to her profession. Women hold a privileged position in the United States when it comes to their opportunities to return to the universities in their middle years. Many are taking advantage of it, either to update former skills or to change fields entirely. Shouldn't we be spreading the message that they are as welcome in electrical engineering as their daughters?

Nearly 24 million women are employed in the U.S. today, comprising over one third of the total work force. More than one half of these are married, which means that of all married women in the country, more than one in three are working.[12] It is a safe bet that all of them must try to match family requirements to job demands as best they can. It is no wonder that the women who work have special needs for good health and a large measure of energy. The number who can manage marriage, motherhood, and graduate training serves to underscore the old adage that the power of women is never to be underestimated. Part-time work in 1968 is limited to certain industrial areas; however, it appears likely that working hours may become shorter in the more populated, more automated world of the immediate future. If so, women will be allowed more choices in balancing career goals and family demands, all to the good of the professions they choose.

Secretaries and beauticians are as likely to work outside their homes as are the mavericks who have families. It does not follow in either case that children are neglected. A psychologist has this to say," . . . we know that women who like their work are also likely to be warm and accepting to their families, that people who accept themselves are also accepting toward other people and toward the institutions of their culture. Interest and enthusiasm seem to expand with practice. In view of the unpredictability of the future, it seems most sensible to encourage the young individual, male or female, to stretch all capacities and horizons as far as possible." My own feeling is that the professional woman who continues to work even when the family does not require her income has a strong need for self-expression that is not met in her role as housewife, and she has a carefully cultivated outlet for this. The untrained woman with the same needs and motivations will very likely devote herself to club work and community service.

It has been well stated that "a good occupation stretches you to meet its demands without demanding the impossible from you. It is one which offers you satisfactions that are not so available to you elsewhere."[13] Occupations have become very important in our culture because so many personal needs are satisfied by them; they give us status and feelings of personal esteem, and they satisfy our desire for self-expression. The women whose special talents and interests lie in the applied science and mathematics category have a real potential for creative contribution and loyal service to the electrical engineering profession. Like men, they will come to an organization with a variety of backgrounds and purposes. All are best motivated by being treated as intelligent adults with full responsibility for the job to be done.

What we can do for her

It is part of our professional responsibility for the continued dynamic growth of our own field to accept women in this light and to encourage those around us to do the same. We can do it directly, by bringing their potential worth as electrical engineers to the attention of students, counselors, parents, teachers, personnel administrators, etc. We will do well to stress the fact that women engineers are real people doing real jobs that other women with the same interests and aptitude could do equally well. Those of us in education have many such opportunities. In addition, we can create a welcoming atmosphere in the engineering classrooms for those girls who have already taken the initial step. We can remind them that they have made a good choice to fit their expected life patterns. Much valuable and rewarding engineering work is done by people who hold the bachelor's degree, in contrast to the Ph.D. degree required in the pure sciences. Choices are available to the engineer regarding the amount of graduate work (s)he will do, and it is important that women realize this. Project engineers and management personnel in government and industry can influence the career choices of entire communities when they view women engineers as colleagues.

Women have a history of positive response to their country's need for their services. The numbers of women entering engineering made significant jumps during both World Wars. This is due in part to the fact that new opportunities suddenly became available, and perhaps in equal part to the fact that women felt needed and wanted. They are needed now. By being both realistic and optimistic about their capability and potential, we have the chance to keep the field of engineering in its place of prominence. At the same time, we will assume a position of leadership in a society that seeks to make intelligent use of its intelligent women.

REFERENCES

1. Barth, E. C., "Engineers in skirts," *Natl. Bus. Women*, Apr. 1962.

2. Peden, I. C., "Women in engineering careers" in *The World of Engineering*, J. R. Whinerry, ed. New York: McGraw-Hill, 1965.

3. Popper, H., "What you should know about women engineers," *Chem. Eng.*, Sept. 11, 1967.

4. "Biennial survey of undergraduate women engineering day students," Society of Women Engineers, New York, N.Y., Winter 1966.

5. Peden, I. C., "An open letter to engineering students regarding women engineers" in *The Trend of Engineering*, University of Washington, College of Engineering, Oct. 1967.

6. Peden, I. C., "The missing half of our technical potential: can we motivate the girls?" *The Mathematics Teacher*, Jan. 1965.

7. Roe, Anne, *Psychology of Occupations*. New York: Wiley, 1956.

8. Hartley, R. E., "Women's roles: how girls see them," *J. Am. Assoc. University Women*, May 1962.

9. Peden, I. C., "Engineering" in *New Careers for Women, 1970–80*, American Association of University Women Foundation, Inc., Washington, D.C., 1966.

10. Saunders, R. M., "Electrical engineering education in 1975," *IEEE Trans. Education*, vol. E-8, pp. 33–37, June–Sept. 1965.

11. "Profile of a woman engineer," pamphlet, Statistics Committee, Society of Women Engineers, May 1963.

12. Harrison, Evelyn, "The working woman: barriers in employment," *Public Administration Rev.*, vol. 24, June 1964.

13. Roe, Anne, "Satisfactions in work," *ASUW J.*, May 1962.

In a symposium that seems to mark a renascence of the science of cybernetics, the roles of man, of machine, and of men and machines together, are scrutinized as purposive or goal-seeking systems

Purposive systems: The edge of knowledge

Nilo Lindgren *Staff Writer*

This report on the First Annual Symposium of the newly formed American Society for Cybernetics describes the papers presented under three broad categories of purposive systems—man, machines, and men and machines together. The discussion ranges from purposive behavior at the level of genetic evolution to the purposive behavior of organisms as complex as modern society. In this period when the growth of technology threatens to overwhelm society, cybernetics is seen as a science offering the hope of solutions and of broader interdisciplinary and international communication.

Cybernetics, the first page of Norbert Wiener's book tells us, is the study of control and communication in animal and machine. The text of the book tells us very much more, more than can be absorbed in the first reading. But the reader certainly understands that cybernetics concerns itself with the most heterogeneous aspects of life and society, and particularly with the intersection of that life and society with technology.

In his preface to the second edition of *Cybernetics*, issued in 1961, Wiener sketches out the scope of the science. "If a new scientific subject has real vitality," he writes, "the center of interest in it must and should shift in the course of years. When I first wrote *Cybernetics*, the chief obstacles which I found in making my point were that the notions of statistical information and control theory were novel and perhaps even shocking to the established attitudes of the time. At present, they have become so familiar as a tool of the communication engineers and of the designers of automatic controls that the chief danger against which I must guard is that the book may seem trite and commonplace. The role of feedback both in engineering design and in biology has come to be well established. The role of information and the technique of measuring and transmitting information constitute a whole discipline for the engineer, for the physiologist, for the psychologist, and for the sociologist. The automata which the first edition . . . barely forecast have come into their own, and the related social dangers against which I warned . . . have risen well above the horizon."[1]

Primarily, the unacquainted may suppose, cybernetics may be regarded as a comprehensive and unifying philosophy, embracing many pragmatic consequences. The supposition is reinforced by a consideration of the breadth of papers presented at the First Annual Symposium of the American Society for Cybernetics held in Gaithersburg, Md., not long ago.[2] The new society, many years in gestation, now emerges, like Gargantua, perhaps holding the keys to a long and active existence. Presiding at its birth was its first president, Dr. Warren S. McCulloch of the Massachusetts Institute of Technology. Vice president of the society is Dr. Lawrence J. Fogel of Decision Sciences, Inc.

It should go without saying that the concerns of this new society are the concerns of many of us.

The edge of knowledge

The theme of the symposium was thoughtfully conceived—"Purposive Systems: The Edge of Knowledge"—and a number of the papers were truly outstanding. For instance, Seymour Papert's conjunction of teachers, circus animal trainers, and the educators of intelligent machines brings color and passion to a subject often hypertrophied by a combination of excessive abstraction, fragmentary example, and sentiments about the dignity of man's intelligence; David Hawkins' appeal for "eolithism" in contemporary human workmanship reveals that poetry is not forbidden in a "technical" conference; and Margaret Mead's staunch strictures to the children of cybernetics remind one that professional societies are really started by strong individuals and that mere organization does not in itself guarantee a long and lively existence.

The incorporation of purposiveness in man-made systems is still a recondite endeavor for the engineer; and the elucidation of the manifold manifestations of purposive behavior, from the level of genetic evolution to the activities of men in technological societies, may leave one feeling alternately dim-witted and exhilarated. In short, the speakers and the audience attempted to illuminate dark territories at "the edge of knowledge." Where they succeeded, they succeeded brilliantly.

The two-day sessions of the symposium, sponsored by the National Science Foundation, were divided appropriately into major sessions entitled "Man as a Purposive System," "Machines as Purposive Systems," and "Men and Machines Together as Purposive Systems," a triad of titles that certainly suggests a kind of classical unity.

Rather than give "equal time" to the many interesting papers, we shall select, largely through the prejudice of personal response, certain ideas that seemed to take wing during the sessions. Sometimes we extract only a sentence, sometimes a paragraph, and sometimes many pages. The standard of choice in each case is by no means uniform.

The major sessions were preceded by a global view of the problems raised by our new kind of technological culture; this presentation, entitled "The Challenge," was given by Dr. Frederick Seitz, president of the National Academy of Sciences.

The whispers of reaction

In his sweeping historical and, in many respects, traditional account of the emergence of a new kind of technological culture, Dr. Seitz traces the descent of man from hunter and food gatherer, to agronomist, to industrialist, to new scientist, to the "Era of the Magic Bullet." Just around the corner, he says, is a period in which essentially all men will live on the basis of a common fund of knowledge. This body of knowledge, he goes on, is now highly operative in the Atlantic Community and along into Eastern Europe. Before the end of the century, it will be effective in most of Asia. And Africa, he surmises, will become part of this common technological culture within a century at the outside.

But the rise of the technological culture brings its own problems owing to the interaction of previously separate problems. There is, for instance, a lack of interest in the kinds of innovation that will help solve the new problems, and Seitz expresses his belief that the great Roman civilization died because of such a lack. Another great issue of our time is the recrudescence of genocide, which has emerged from a 2000-year sleep, a seeming concomitant to the rise of technology.

The danger of a failure of interest in innovation in the United States seems to Seitz quite real, for in his rather general peroration the image emerges suddenly garbed in detail. It's a trend, he notes, "about which Teller speaks these days, the tendency to downgrade technicians." That term, Seitz goes on, "has not yet been used in a degrading sense for scientists, but on Capitol Hill, one hears more than faint whispers of it. Senator Pastore attacked the great nuclear accelerator, saying that it is just something for the academic scientists," implying that it is not important and, to Seitz, suggesting that our society may indeed go the way of Rome. Given "not a generation, but ten generations" for that kind of reaction to evolve, we will see "something quite different from what we have now." Echoes of such a reaction are reported from the recent meeting of the American Physical Society in Chicago, where the "Cassandras" spoke of a "revulsion against science" throughout all American society.[3]

Against such a reaction, and the problems that threaten to overwhelm us through their complexity, Seitz balances the hope of cybernetics, which could be a means of unifying many separate disciplines. Accompanied by a continuing growth in computer technology, even a "small professional cadre" of cyberneticians, Seitz speculates, could handle vastly more complex problems and give renewed vitality to our human society.

Cybernetics and society

One must note that throughout the symposium, there ran like a natural countertheme to that of purposive systems a deep concern about what is happening within our society, what technology is doing to it. This concern, expressed in many ways, reminds one of a statement made by David T. Bazelon at another conference held in Washington about the same time. He said, "*My main point is that technology does not change society; it destroys it.*" By this he meant "that the foremost precondition of the wonderful wealth–machine, the industrial system, is the destruction of the previously existing society."[4]

The "Grand Lady of American Science," Dr. Margaret Mead of The American Museum of Natural History, also rings this theme in her "Commentary." When she was asked to speak at this cybernetics symposium, she had this impression: "I'm not sure he said trouble, but his voice said *troubles*, and I wasn't quite sure if it were troubles of cybernetics or of the world that he was referring to."

A commanding and humane figure, Dr. Mead was a member of the original Josiah Macy group at the beginnings of cybernetics, back in 1943 or so. Considering the ups and downs the cybernetic movement has undergone in the scientific community's favor, the kind of perspective Dr. Mead gives us is valuable.

In an amusing preamble, Dr. Mead begins by telling us about the things she is *not* going to talk about. She is not going to talk about the influence of computers, the electronic revolution, implosion versus explosion, the decay of books, the importance of dress (although she cannot help noting that dress has succeeded the mimeographing machine as a form of communication among the young); matters naturally of interest to the anthropologist concerned with what is happening in our culture and who stands outside of it to a certain extent looking at it as a whole. Rather, she identifies and limits her interest, for the purposes of this "not initial but first fully organized" symposium, to the question of what cybernetics is going to do in our society. How, she asks, can cybernetics help us to handle the increasing complexity?

Specifically, she undertakes to talk about "the significance, as a form of cross-disciplinary thought, of the development first called feedback, then teleological mechanisms, then cybernetics." One of the original motivations of those who, like herself, worked in the first meetings on cybernetics, Dr. Mead says, was that cybernetics was a way of looking at things, a "kind of language" with which many disciplines could communicate. It constituted a "language sufficiently sophisticated so that it could be used at the top of our attempts to solve human problems, and sufficiently abstract so that it was possible to cross the boundaries of the disciplines. I think this can hardly be overestimated although we haven't used it as extensively as we might. When we developed it, we thought that we would go on into interdisciplinary research using this language. Instead, on the whole, things fragmented. Wiener wrote *Cybernetics* and it became a book for certain kinds of intellectuals for a time. It looked as if it were becoming very much a part of the vocabulary of intellectuals, but then it decayed. It had had its life."

Now, Dr. Mead goes on, young people do not understand the meaning of cybernetics as it was interpreted in the beginning. "There was a period about 20 years ago," she muses, "when both the United States and the Soviet Union thought there was nothing else on this planet but the other." Each was trying to penetrate the other's secrets and "hypothesizing a good deal of systematic behavior that wasn't true. . . . Even as late as 1961, it was possible to think of the United States and the Soviet Union as two systems coupled together by mutual suspicion, passionate interests, and intermittent surveillance." In that situation, Dr. Mead proposed that cybernetics might be used as a cross-vocabulary between the two nations, because she saw it as being free of ideology. Through it, for instance, she thought it might be possible to discuss the differences between the economic and political systems, and find the areas in which the two systems wanted to agree, "which was oftener than anyone wanted to admit."

As we know, this did not happen. Dr. Mead traces some of the factors and fears that defeated that particular hope in the years before 1961. But now, once again, she sees the possibility of bypassing ideological constraints through the teaching of "general systems theory" around the world, for it provides a set of tools for thinking about complex systems. "I think," she says, "as the world scene broadens that the continuing possibility of using cybernetics or general systems theory in this broader sense as a form not only of interdisciplinary communication, which was the way we thought of it then, but as international communication in a world of increasing national specialization in science, is something we ought seriously to consider."

Furthermore, she urges, in our societal attempts to handle rationally large systems, such as the growth of metropolitan areas or the interrelations between different levels of government, we should recognize that their dynamics are a concern of cybernetics. We have had thus far, in such extremely complex systems problems, "far too much linear planning and too little lateral planning." And she cites a number of examples, such as the massive Northeast Blackout, in which there has been a failure to think out adequately the nature of large systems.

And her anthropological outlook, famed for its analysis of small, intimately known living communities, where usually "there are a lot more mosquitoes," becomes more apparent: "At present, I think that the introduction of automated systems and circular systems of one sort or another into a society that does not understand them, that fears them, and that is willing to explain almost anything in terms of them . . . unless it is looked at very sharply will create an atmosphere in which it is impossible to take the next step in the effective societal use of the skills that we do have." She here shares some ground with Dr. Seitz.

Warren S. McCulloch

"In many instances," Dr. Mead adds, "we are building systems where the use of human judgment and the use of automatic decision making are put in a very bad mix, and we aren't considering sufficiently the time relationships between the two." Moreover, "there has been a progressive diminution in the willingness to look at the human component in a large variety of complex systems . . . and we are not considering seriously the education of those who are going to have to become the human computers in man-machine systems, those who are going to have to make the decisions about them."

With such and many other warnings about the consequences of living in a world with the diffuseness of modern scientific endeavor, Dr. Mead comes at last, also, to some hard questions addressed to the founding fathers and new members of the Cybernetics Society. *What* is being founded, and why? How is the society to keep from getting old? How is the vocabulary to be kept sufficiently current? Does the society wish to die in ten years? "This is one of the more intelligent things that lots of organizations do now. They set a terminal date and say the thing we're doing won't be worth doing in this state ten years from now." In sum, she seems to be asking, "How can you insure your continual renewal?" It is a question that every social organism must continually ask.

Man as a purposive system

It is appropriate to begin the description of the technical sessions with a definition of purposive behavior as Dr. Alexander Fraser, a geneticist from the University of Cincinnati, did in the opening paper, "The Evolution of Purposive Behavior." Fraser distinguishes the evolution of two forms of purposive behavior.

"One is the evolution of a pattern of behavior, just directed towards achievement of an intrinsic purpose. . . . Natural selection sets the intrinsic purpose, and the organism adapts to fit that purpose."

The second is "the evolution of the ability to define purpose and to elaborate behavior to achieve it." But, in terms of evolution, the ability to define a purpose has an intrinsic purpose hidden behind it, namely, survival.

Fraser reviews how genetic modeling has been done on computer. In elementary biology courses, he notes, you deal with one or two genes, but the real-life geneticist must deal with anywhere from 30 to 60 genes at a time, which means a fantastic level of complexity. It means, for instance, "examining changes in the three-to-the-30th space."

The major evolutionary force in the real world, says Fraser, is the survival of the mediocre. Deviant individuals of a species tend to be discarded—those who are too big, or are too small. "Being clever is possibly highly dangerous."

Yet, he points out, from an evolutionary point of view, a random event can set a whole population on a track. Fraser ends by making comparisons with the evolution of machines. "When we're talking about purpose in evolution," he says, "there is a tremendous tendency to look around at the superlative machinery of living things and consider these as being put together all in one block, whereas in actual fact, it is far more likely that evolution blundered along making 50 percent compromises that gradually built up on top of each other until they got to the present system. . . . I must admit to seeing direct analogies between genetic learning and learning in the individual sense. There are eminently inefficient compromises. . . . One of the first mistakes we make in trying to get a machine to evolve a language is in expecting it to evolve a language that Shakespeare could use. I would be satisfied if we could get a machine to do anything at all to begin with. After all, the very first viruses were almost certainly fantastically inefficient in their energy use compared to later viruses, which are fantastically efficient." The same kind of idea about machine learning is affirmed in the paper by Seymour Papert.

Fraser's discussion of the evolution of purposive behavior was followed by "The Neurophysiology of Purposive Behavior," presented by Dr. Ralph W. Gerard of the University of California, Irvine. Dr. Gerard notes that he had come to very much the same kind of conclusion from a neurophysiological point of view as Fraser had reached from a genetical point of view, "first in being willing to use the term learning for genetic and evolutionary processes as well as those within the individual and, second, for regarding evolutionary or genetic learning as . . . a matter of learning to learn."

Although Gerard discusses purposive behavior in living systems, he first offers an example of such behavior in a nonliving system, which the reader might find a useful steppingstone to more abstract considerations. The example is Shannon's mechanical mouse, which finds its way out of a checkerboard maze. It was demonstrated at an early cybernetics meeting. The mouse ran into the walls all along the maze, but finally came out at the end, after innumerable blunders. Put back at the starting point, it went directly, without an error, through the correct squares. Its mechanism was set simply so that the direction in which the mouse left a given square for the last time was the one that it remembered. This was, says Gerard, simple, elegant learning in the best sense. The mouse was automatically programmed for completely purposeful and correct behavior.

However, in "living systems, you find a richer and more expensive array of purposeful behaviors that clearly have nothing at all to do with the nervous system." For instance, in a primitive living system like a seed in the ground, the stem goes up and the roots go down, a highly purposive behavior, with no nervous system remotely involved in it.

Making his way through a series of interesting examples of purposive behavior in animals, Gerard closes in on his subject of purposive behavior in the central nervous system, and the mechanisms underlying them. "Fifty years of neurophysiology," he says, "have moved us from the telephone exchange model (fixed in time, space, and number, and able to explain repetitive, invariable behavior) to a dynamic, spontaneous, modifiable model (free in time, space, and number, and able to explain variable, creative, and integrated behavior)." Instead of thinking of neurons as sitting inertly in the brain waiting until a message comes along to activate them, neurophysiologists have come to the recognition that neurons are physiologically active essentially all the time, going through rhythmic discharges, fluctuations in threshold, progressively summing impulses coming in from different channels, and so forth.

Further evidence, implicit for a long time, but now precisely worked out by many investigators, shows that "the flow of messages along particular patterns of neurons leaves behind some kind of facilitation of future movement of patterns, of similar patterns, along similar neurons." And that, Gerard notes, is the crux of the learning experience.

Although we cannot follow his arguments here, Gerard cites experiments that have led him to postulate what he calls a physiological neuron reserve, which, when it has been cut down, and when the anxiety level of the organism has risen too high, leads to a kind of stenotic or nonadaptive behavior on the part of individuals and even of nations.

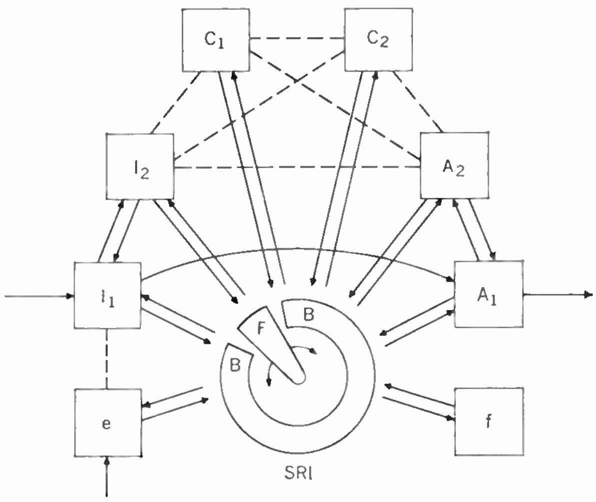
Models of information processing within man

In his recently translated book, *Modeling of Thinking and the Mind*, Dr. N. M. Amosov, Deputy of the Supreme Soviet and head of the Biological Cybernetics Department at the Institute of Cybernetics, Kiev, states that the "creation of thinking and feeling automata has already passed out of the sphere of the fantastic and has completely real foundations. This is a scientific problem, and its psychological and social aspects must be examined with all seriousness and responsibility."[5]

In the scientific enclaves far west of Kiev, there is probably no deep quarrel with the concept of automata that think, but the concept that automata will also "feel" may be met with some doubts. Yet, it is clear from his book and from his presentation at this conference, "Models of Information Processing Within the Man," that Amosov incorporates programs of feeling, consciousness, and even creativity in the highly elaborated models of human thinking and behavior he hypothesizes.

Essentially what Amosov offers is a complex chain of heuristic computer programs, based on principles of simulation and information processing, and assembled with some insight into the nature of human cognitive activities.

FIGURE 1. Model of a system of reinforcement-inhibition (SRI) proposed by Soviet cybernetician Nicholai M. Amosov. Shown here are: the reinforcement of a model (F); the inhibition of the other models (B); and four different types of cortical models. The boxes labeled I_1 and I_2 are the hierarchical models of the meaning of the surrounding world; A_1 and A_2 are hierarchical models of outwardly directed actions; C_1 and C_2 are models for programs of consciousness; and e and f are models of emotions.



These programs are built up into hierarchies of information processing that have participating parallel subprograms. An artificial intelligence with all the attributes of human personality, Amosov claims, can be created on such a foundation.

He identifies six types of programs that determine human behavior:

1. A program for bodily feelings, perception, and processing of information from the body. Feelings, he says, are cortical reflections of the excitation of subcortical centers, representing instincts, such as self-preservation, reproduction, and programs of social behavior. The interconnection of such "subjective" elements with other programs representing consciousness allows such factors to influence conscious decision making.

2. Programs for the perception of outer influences from the environment and for the recognition of the meaning of environmental qualities.

3. A program for action, that is, for the imparting of both energy and information to the environment.

4. A program for speech capabilities.

5. A program for consciousness.

6. A program for creativity and labor.

Amosov offers block diagrams showing how these kinds of heuristic programs are interconnected. To indicate the character of these programs, Fig. 1 shows Amosov's method of solving the problem of dominance. Amosov calls this a system of reinforcement-inhibition (SRI). What it does is to select and reinforce only one (the most active) of the different cortical models (feelings, consciousness, actions, meaning of the environment) at any given moment, during which time the other models are inhibited. After a short time, the connection of the model with the SRI becomes "tired," and the reinforcement turns to one of the other models displaying the most activity. In effect, what Amosov gives us here is an overall model of how the attention of the hypothetical organism shifts. At any time, any of the models—images of the outer world, feelings, actions, desires, etc.—may exist (be

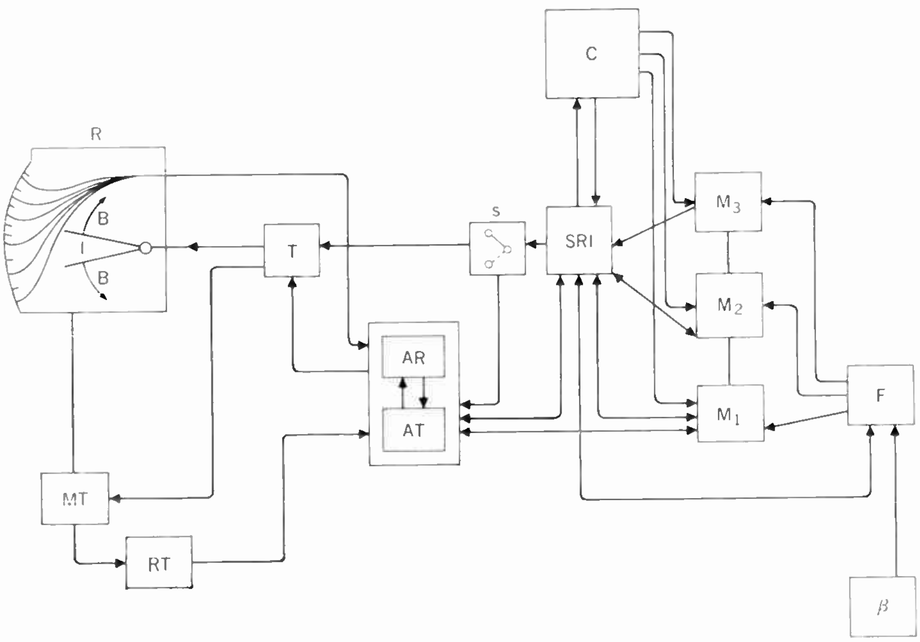


FIGURE 2. Models for the perception of environmental effects, selection, and information processing show how the SRI of Fig. 1 is incorporated. The symbols show: the main receptor (R); the muscular attenuation or orientation of the automaton to the outer world (MT); the receptor of attenuation (RT); the attentuation (T); the reinforcement of a portion of the receptor neurons (I); the inhibition of the other neurons (B); the analyzer of the main receptor (AR); the analyzer of attentuation (AT); the relay of SRI from the receptor to the analyzer (s); models of consciousness (C); the influence of feelings (F); the body (β); and hierarchical models of the environment (M_1, M_2, M_3).

connected to) the sphere of consciousness, and, as Amosov says, the model reinforced at a given moment constitutes a *thought*. To go just a bit further, Fig. 2 shows Amosov's program for perception and the processing of information from the environment.

To test his hypotheses, Amosov has simulated a model of purposive behavior in an automaton consisting principally of two major blocks, one for action, the other for information processing during the action. The automaton was made to "walk" to a planned destination while interpreting objects of its surrounding world as "forest, beast, or food," and evidently taking appropriate action as it went along. (Although the map of its route shows that the automaton certainly took steps to avoid the beasts, it seems to amble quite happily through the forest in search of the next supply of food.) A chart of the levels of activity in the automaton's "consciousness" and "subconsciousness" shows much more going on in the subconscious. Shades of Freud!

Amosov concludes that his hypothesized model could serve in the study of basic mechanisms of nervous activity, both normal and abnormal.

Machines as purposive systems

When one considers, on the scale of natural evolution, the relatively short time that men have been attempting to build machines that are purposive, it is quite understandable that there is still much less to say about such machines than about the living systems they will presumably mimic. Yet, the attempt to make purposive machines has already exercised a deep influence on the way man sees himself. The endeavor has given him new languages and a new perspective. When Amosov discusses a program for feelings in an automaton, for instance, one is obliged to re-examine one's own conception of the nature of feelings.

The issue is taken up, too, in another way, in a coruscating presentation, "Why Machines Can't Think," by Dr. Seymour Papert of M.I.T. Dr. Papert's point is that machines can't think as yet, or at least can't think better than they now can, because we haven't given them the possibility to be intelligent. Most "fully fledged proofs" about what machines can or cannot do, he notes, "depend in the end on what people have done rather than on what is true or false of machines. Machines can't think now nor will they ever as long as people don't give them a chance."

In what way, Papert asks, have people not given them a chance? One way, he notes, is that "we have grossly neglected the question of education of machines." That is, on the question of how to teach machines, he makes this distinction: "Everybody knows that a lot of work has been done on so-called learning machines; but to study how machines learn is not the same thing as thinking about how you would go about *teaching* a machine. As Warren McCulloch keeps telling us, although we sometimes have difficulty hearing him, the situation is a triadic relation between teacher, student, and what is being taught." It is by no means a symmetric relation. On the science of education, Papert quotes his teacher Piaget: "Before any teacher goes into a school, he ought to spend three months in a circus with the animal trainers because, in the circus, if the tiger won't learn the responsibility is with the trainer." And Papert pointedly asks, "If the machine won't learn with whom is the responsibility?"

Papert then traces the history of the efforts in teaching machines how to play chess, long considered a "paradigm for the study of machine performance in an area that seems intellectually difficult for humans but that is sufficiently circumscribed," so that the whole gamut of knowledge, language, etc., need not be taken into account. In his discussion, he cites a chess-playing program, written by R. Greenblatt at M.I.T., that differs materially from earlier programs and that has a bearing on the question of education. This program allows the player to make moves, to get replies from the machine within a few minutes, to retract moves and to make alternative ones, to ask the machine to print out the set of moves it considered. In short, Papert says, "sitting at the console and interacting with the machine, you can do something that is reasonably describable as exploring the intellectual personality of the machine as a chess player. It is an interaction more akin to the art of the psychiatrist or the real teacher." Thus, the player, bit by bit, explores the weaknesses of the machine, and improves its playing powers by correcting bugs in the algorithms it has been given. The program, which has played about a thousand games by now, has been going from success to success in gradually increasing its chess score. The moral, which Papert draws after a series of engaging moves, is that when we deprive the machines of the kinds of possibilities that we use ourselves and "then complain that machines cannot be intelligent, we're like the people who put the pigs in the filthy sties and call them dirty animals."

J. A. Haddad of International Business Machines Corporation took up the task of identifying trends in hardware development that would facilitate the design of purposive systems. In his presentation, "Hardware for Purposive Systems," he noted that it is really necessary to consider the hardware and software in conjunction. He pointed out, also, that "in an information system the goal may be imprecisely defined and the purpose of the system then is to help define the goal. . . . Once the system defines more clearly the goal from the information it has, it can then seek to arrive at the goal. In so doing, it may generate further information that causes it to refine further its definition of the goal. In some cases, this process does not converge, or does not converge fast enough. In such cases, the system should automatically give up seeking that particular goal, so inform its operator, and seek some other goal." Thus, Haddad arrives at the deduction "that computer hardware *and* computer software are necessary in combination in order to have a system that is purposive."

He traces ever-higher purposive behavior in an extremely complex environment, in which there are many decision points in the process and thus many feedback loops of a very subtle nature, demonstrating how a user and a computer are both goal-seeking simultaneously in a symbiotic relationship.

Haddad sees a significant development in the emerging practicality and utility of associative memory in central processing units. Such associative memories, which will contain in the stored data both the argument and the function, will be essential in purposive systems.

Even more powerful for goal-seeking systems is the development of control memory. Haddad notes that "there is no basic reason why one cannot design a control memory with a . . . technology that would allow the computer itself to alter the information stored in the control memory. Thus we would have a computer that could alter its

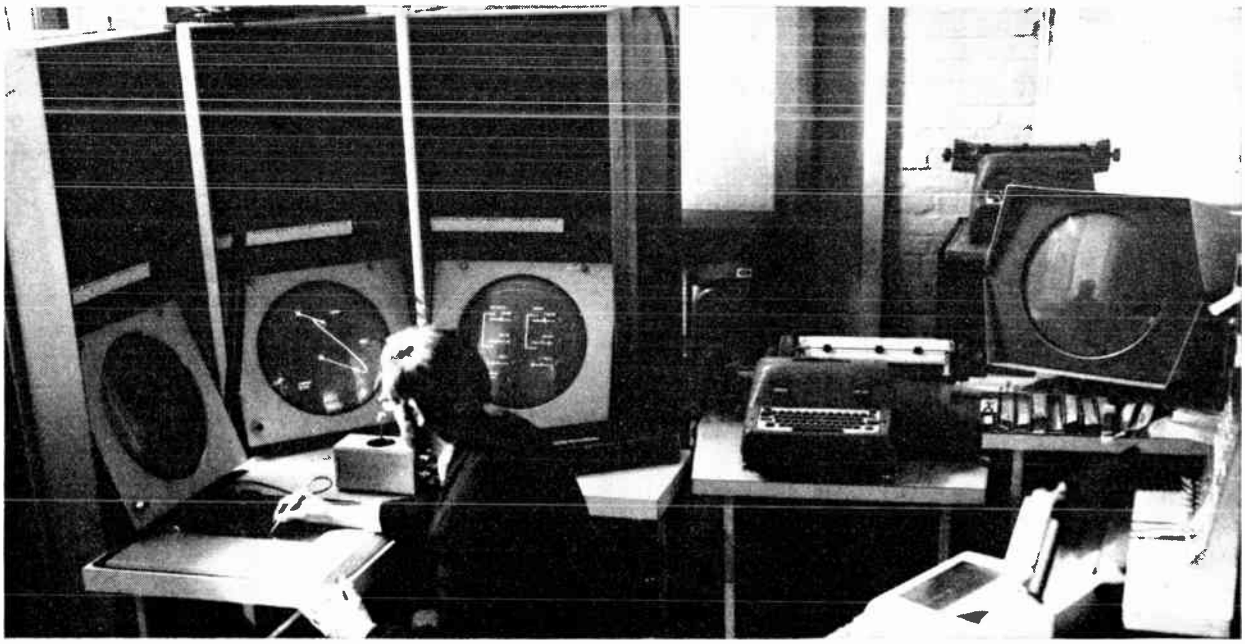


FIGURE 3. View of the multichanneled computer input-output display console being developed at Harvard Computation Laboratory. Graduate student Ted Lee manipulates parametric curves on a PDP-1 computer.

own character as required. To my knowledge very little conceptual work has been done in thinking through the implications of this extremely powerful possibility. The potentials are so staggering and deep that the poor harried souls who are trying to understand the classical computer as we now know it wish this idea would go away."

When one even begins to conceive such a notion as the "classical computer," it becomes clear that we are encroaching on new territories. "It is a fact," says Haddad, "that computer hardware today is capable of purposive behavior when combined with software."

In this same session, Dr. Saul Amarel of the Radio Corporation of America presented "Problems of Representation in Artificial Intelligence," in which he discusses the nature of goals in problem-solving environments. Using the famous "missionaries and cannibals" problem as an example, he generalizes his solution to point out that it is of value to search for groupings within the problem space, that is, in the course of transformations to identify variables that are conserved and to recognize analogies in the different kinds of formulations of problem structures.

More general questions on the theories and interpretation of cybernetics were raised by Herbert Anschütz, Ministry of Defense, West German Government, in "Prospects for the Development of the Psychocybernetics of Intelligent Behavior," that is, on the question of making machines that think as we do.

Anschütz notes that although a number of scientific "streams have merged into what is today called cybernetics, the lack of a comprehensive mathematical theory leads into a peculiar divergence of thoughts about what cybernetics is." He mentions, for instance, that cyberneticians use information theory, the theory of control, automaton theory, and other basic mathematics, but that "many cyberneticians, including Norbert Weiner, had a strong feeling that these three theories might really be only different cases of a single unique theory of cybernetics." Possibly, Anschütz continues, the unity of these theories is masked by their usual descriptions. That is, Shannon provides an essentially statistical theory of information; automata theory has thus far been primarily deterministic; and noise within control systems has been viewed as unavoidable disturbance.

Anschütz suggested "some facts that when brought together might lead to a single theory of cybernetics which includes all aspects of human thinking." He goes on to describe a new measure called "eneidy" (Greek for "to seem") that is analogous to the thermodynamic measure called energy. He applies this new measure in inanimate systems and suggests that it will help to quantify the interaction of social systems.

Men and machines together as purposive systems

Opening the third major session of the cybernetics conference was Dr. Ivan Sutherland of Harvard University, of whom Dr. Licklider said, he "talks about highly crystallized things in a highly organized manner and, you know, there is nothing better" than that.

Sutherland's subject, "Facilitating the Man–Machine Interface," deals with the research trend toward three-dimensional equipment, and particularly with the work in his own group at the Harvard Cruft Laboratory. The group is developing a perspective-generating device that will display images in true perspective rather than in orthographic projection. Sutherland describes the mathematics involved in producing curves in *n*-space, citing examples, and discusses the "hidden line" problem, namely, the difficulty of eliminating the parts in the 3-D view that wouldn't be seen if the object were solid. He also brought along a fascinating and unexpectedly amusing film showing a display system that actually analyzes simple mathematical expressions directly, and that replaces crude lightpen handwritten characters with its own machine version.

The display equipment under development at Harvard is shown in Fig. 3. The console includes a typewriter, a teletypewriter, four or five oscilloscopes, a RAND Tablet, a printer, five tape recorders, a drum backing up six blocks of core memory, and so on. The facility evidently will incorporate sophisticated three-dimensional display equipment evolving out of the research that Sutherland describes.

One of the keenest problems in man–machine interaction has to do with the languages that will be used in such interaction. Addressing himself to "The Future of Man–Machine Language Systems," Dr. Yehoshua Bar-Hillel, of Hebrew University, Jerusalem, notes that the "amount of intelligent conversation possible today between man and computer is extremely restricted, both as to form and substance. . . in spite of the not inconsiderable publicity given to some such conversations."

Bar-Hillel has previously proffered his convictions that machine mastery of human language would not come about, and he does here again in the form of three criteria for machine intelligence. The machine, he says, should have at least the faculty of (1) manipulating natural language at the level of a high school graduate, (2) background knowledge at about the same level, and (3) reasoning and computing abilities of the same level. None of these prerequisites is fulfilled at present, says Bar-Hillel, and, more significantly, there is nothing in view that should make people believe that simply with hard work and a lot of money such prerequisites could be obtained in the foreseeable future. The required effort, he says, would be incomparably greater than that required for putting man on Venus. Even with "the almost incredible achievements of Chomsky and his associates in providing us for the first time in history with rudiments of a linguistic theory worthy of the name," Bar-Hillel sees no real basis for intelligent man–machine conversation, nor does he see any long-range prospect for the construction of such a system. "The only rational way of making any progress," he says, "lies in the radical lowering of our aims."

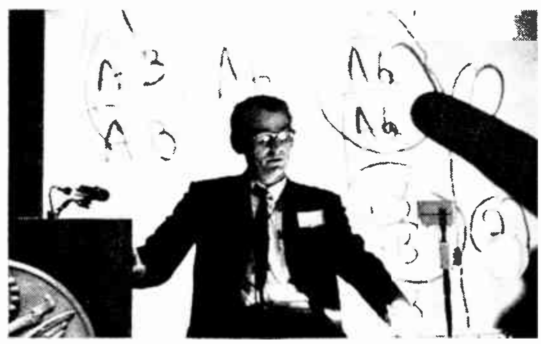
In this same session, Dr. Talcott Parsons, president of the American Academy of Arts and Sciences, in a presentation entitled "Facilitating Technological Innovation in Society," devotes himself more directly to social concerns, as does Dr. Emmanuel Mesthene of Harvard in an earlier presentation, "Social Organisms as Purposive Systems." Mesthene expresses the conviction that "the most fundamental *political* task of a technological world . . . is that of systematizing and institutionalizing the social expectation of change." And Parsons sees "the university on the way to becoming the central structure of the coming phase of modern society."

Man–machine interaction

"Most creative work is not," says J. C. R. Licklider, "solving already formulated problems, it's formulating them." And this is probably one of the things that Licklider does best. In the arena of man–machine symbiosis, he blocks out for his scientific and technical confreres the next big research problems that lie ahead, and he tells why they need to be solved and where they are likely to lead. It is as if he were living ten years ahead; he gives an impetus to his vision with a cajoling humor. Given a real taste of what does lie ahead in Sutherland's laboratory at Harvard, he exclaims, "Oh, come on, technology, hurry up! Or if it's not technology that's holding us back, economics, get out of my way!"

Margaret Mead

Alexander Fraser



The question he takes up is how interaction between men and computers can lead to the understanding of things that are not well understood, and how such interaction can sharpen the formulation of problems that are not well formulated.

In the first place, he says, this kind of work is not what has earned computers their reputation and established them as absolute musts in business and industry and in the solution of scientific problems. Their major use over the past two decades has been in the lightning-fast, extremely accurate solution of already formulated problems. However, Licklider continues, as far as purposive systems are concerned, we are not interested in this final stage of problem solving, which isn't really problem solving anyway, but merely a kind of numerical evaluation process.

The kind of man–machine interaction that Licklider is

W. S. McCulloch, N. M. Amosov, and F. Seitz

talking about does not have a long history. Back in 1959 or so (Licklider's germinal paper, "Man–Machine Symbiosis," was published in 1960[6]), the concept was a gleam in the eye, and, as Licklider points out, was disregarded or thought to be unimportant by almost everyone. But by now, things have changed, and there are many people in the United States at least who have had real experience in interaction with computers. Unfortunately, Licklider notes, most of the experience has been of a very primitive and limited kind, consisting mainly of interaction through a typewriter or teletypewriter.

"If I were to try to characterize the present situation," he continues, "I'd say that it is one in which people have started to appreciate how very much better it is to have interaction facilitated by computer graphics and by direct access to watts of computer power than to have the metered-out conversational interaction through a typewriter. We are in a state of real confrontation between a potential that makes us impatient to plunge ahead and economic restrictions that prevent us from giving a few hundred thousand people the kinds of consoles being developed at Harvard." That kind of console constitutes, Licklider says, a "privileged interaction."

For problem solving in the creative sense stipulated, what kind of performance should the new interaction equipment provide? Licklider identifies a number of factors or facilities that do not yet exist, but which could be made.

The first thing he mentions in direct interaction is that the computer response time need not be absolutely immediate. What is crucial, and what makes on-line computing different in quality from off-line batch processing, is that the response be in accord with the span of thought about the topic. For some elements that are low in the hierarchy of the overall problem, the response should indeed be very rapid, as, say, in the computer's recognition of single typed characters. At higher levels, the response may be much slower—for instance, one-day turn around may not be obstructive to a problem that requires days or weeks of concerted effort to think through.

His second factor is the nearly total absence of a macroscopic representation of a problem in existing computer systems, an absence most marked in programming and debugging programs. Many systems show the content of a particular register of the computer memory, but for complex problems with many parts, what is really needed is a macroscopic view, so that parts of the problem can be expanded, as though with a magnifying glass, while the spatial array of the entire problem space remains in view. It is frustrating, Licklider points out, to move in the problem space by typing, either symbolically or numerically, and seeing, as it were, one by one the little electrons of the problem when you are really interested in seeing how such small elements interact.

The third item is what Licklider calls revocability, namely, the capacity for making computer informational processes reversible. In the formulation of problems in a trial-and-error way, in which any hypothesis may be incorrect, it is important to be able to back up and then move on along a different track. This characteristic is especially important for big problems, for which, as it is now, the user may be obliged to start a problem solution completely afresh with the loss of minutes and even hours. Although it might not be feasible to record the entire history of a complex problem, there ought to be at least a spatial separation of problem-solving events at, say, the junctions of subroutines.

Contextual determination is the fourth factor. Most computer languages, Licklider notes, are "context-free," and they make use of the problem context in only the most simple-minded ways. That is, there are frustrating language problems involved in shifting among different programs in the software system. What is needed, evidently, is something similar to what obtains between people, who, in their interaction, sense through shifts in vocabulary that different aspects or levels of a problem are being discussed. Dialogue between people is not context-free. In the computer, says Licklider, the availability of resources ought to be controlled through selection of vocabulary, not the vocabulary controlled through selection of subprograms.

Licklider mentions other needs, such as the "concept of metasimplicity" in problems of real complexity, and the use of field-oriented languages, and how certain input–output devices further complicate already complicated problems.

More germane here, we think, is his conclusion that the realization of real man–computer interaction depends upon a huge amount of software preparation that "only very bright people" will be able to bring about. It will mean a situation for a few years "that looks terribly inbred," as the best on-line computer systems are used for the development of software for on-line computer systems. Now ordinarily, Licklider says, you shouldn't admit that; if you're in a situation that is trying to finance itself, you don't admit that you aren't going to deliver tomorrow. But, he concludes, it will be a big job, taking several years of many people's efforts, to build a base of procedures and a base of data that will support creative, on-line interaction.

The best hope for building such a base, Licklider sees as the concept of a network of on-line interactive, multiple-access computers. By utilizing telecommunication facilities to net together such systems as Project MAC at M.I.T., the time-sharing systems at System Development Corporation and at Carnegie Tech. and so on, the creative, academic, intellectual types of people, although geographically separated, could communicate and work together in the same system.

Such a network, Licklider warns, will not guarantee coherence in the product, the software base, but the lack of such a network will certainly guarantee its incoherence. And, more than anything else, what is now needed is a coherent software base usable by anybody, no matter where he is or what particular computer facilities he has.

The nature of purpose

It was most satisfying, at least to this observer, to hear the final paper, by Dr. David Hawkins of the University of Colorado, who took as his task a discussion of "The Nature of Purpose." In a quietly amusing, lovely, and even poetic meandering (which he himself characterized as being "eolithic"), Hawkins separated out two essential principles of human workmanship—one called the "principle of design" and the other, "by a slightly bastard etymology, the principle of eolithism"—which he draws from the work of the engineer-novelist, Hans Storm.[7] From a contrast of these two principles, Hawkins makes his way toward a Sartrian existential viewpoint: "You know that at any particular moment you are making a choice of being what you are." And from that, he makes the connection between choice and purpose. "If we can get away from thinking of choice as a choice among antecedently given alternatives, and think of choice as a constructive process, then we are at the beginning of some wisdom" about the nature of purpose. From that point, Hawkins moves easily to the concern of the field of cybernetics with the setting of goals, a concern that certainly relates to the way men design their society and their machines. It is a concern, then, that does not belong alone to men who call themselves cyberneticians. Or, as Anschütz suggested, "while this conference is in progress, everyone in this room is a cybernetist. But when the conference is over, each person is once again a biologist, a chemist, an engineer . . ."

Hawkins begins his thesis with the observation that throughout the conference, the theme of "purpose" has been taken in the same sense that purposes are taken when they are by implication already well defined. But Hawkins wants to look at the more active sense of "to purpose," and so he resorts to the images of eolithism contrasted with the standard notion of design. In the standard notion of design, say of a bridge, a building, or an electronic system, the designer, Hawkins says, must first know approximately what he wants and how it is to be used. He next chooses his materials of known, uniform properties. The certainty of the designer's objective and the uniformity of the materials are essential to the whole process, for they affect not only the physical aspect of the result (i.e., the geometry of the architect's design), but they deeply affect the mental discipline involved in the design process as well. With these "givens"—the clear objective and the known materials—the design begins to emerge, is checked and extended through the known algorithms of mathematics, and iterated in detail until the whole becomes realistic, "making contact with the existing world."

At that point, Hawkins goes on, the switch is made from thought construction to actual physical construction. In the contact with the material world, the direction of the process is reversed as the structure is extended gradually until it embodies the objective from which the whole process started.

This traditional principle of design is traced from engineering textbooks and from practical engineering experience. It is the standard; it forms a recognizable description of much engineering work. It suffices to say that the process results in products that have "a certain internal consistency, a certain finish, an orderly and uniform style," and that are "unchildlike," a description that characterizes the modern manufacturing method. The power and advantage of the method is that it leads to high productivity; its disadvantage is that it leads to the blight of "sameness" in the environment of the highly industrialized society.

But Hawkins, and his predecessor, Storm, challenge the assumption that this standard principle of design in our society is basic and universal, that it is an ideal by which we can measure all craftsmanship from its most primitive and blundering beginnings, that it is the ideal to which all craftsmen are committed, whether they know it or not, or

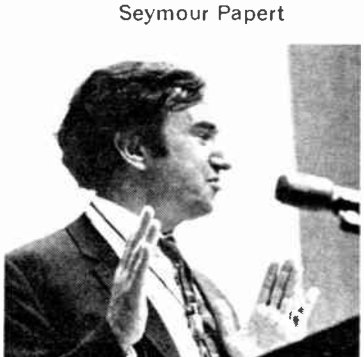
Herbert Anschütz

David Hawkins

J. C. R. Licklider

Heinz von Foerster

Seymour Papert



whether they like it or not. In this challenge, Hawkins continues, "Storm, a professional designer himself, puts forth an alternative, a wholly different principle of workmanship, one . . . more distinctly human than the principle of design . . . for which he borrows the term eolithism. An eolith is literally a piece of junk remaining from the Stone Age." It is something found in the garbage heap, perhaps rescued from an ancient buried garbage heap.

The human eolithic design process has a distinct relation to animal craftsmanship. For instance, nest building by birds seems to be "the accidental by-product of a number of specific little programs," which Storm calls tropisms, "of congenital behavior released by the occasion such as picking up twigs and dropping them with what would appear to be a studied casualness and no evidence of interest in the final product which happens as though by a conspiracy of nature to assemble itself in the end." The emergence of a definite character in the building of cities, which are the accretion of the "accidental" or aleatory acts of multitudes of individuals, might be regarded as an eolithic product of the whole society.

However, Hawkins notes, there are not many examples of eolithic craftsmanship among humans. He cites the collectors of stamps, crockery, automobile parts, but generally, he says, it is a relatively rare style among us. In our society, the eolithic pattern has less esteem and social status, "while most men occupy themselves with war, literature, business or odious routine for wages." The attention to "found objects" as pieces of art, for instance, certainly a form of eolithism, still tends to be regarded as avant-garde activity that perhaps only artists indulge in.

How are eoliths used? They are "stones picked up and used by man and even fashioned a little for his use." The important point is that the eolith is already accidentally adapted to some end, and more essentially, "strongly suggestive of the end." "We may imagine," says Hawkins, "the person whom the anthropologists describe so formidably by the name of man, strolling along in a stone field, fed, contented, thinking preferably about nothing at all (for these are the conditions favorable to the art), when his eye lights perchance on a stone just possibly suitable for a spearhead. That instant, the project, the very idea of the spear, originates. The stone is picked up, the spear is in manufacture." The shaft and the thongs to hold the head in place remain in the background during the shaping work, just as the knotty problems of man–machine interfaces remained subsidiary while the hard core of the computer was being fashioned and assembled. If the spearhead fashioning goes wrong, the piece of stone might be diverted to another use that suggests itself.

Although each design process may include a few ingredients of the other, the contrast is sharp. The designing workman knows what he wants before the design begins, and he must decide on the material, whereas the eolithic fashioner must have a continually open mind about materials and be adaptable about the ends he wants. Where the designing workman is limited by the fixity of his goals and the need to eliminate variety from the means and materials (which are thereby reduced in their significance for other ends), the eolithic fashioner, confronted by an eolith that defies the use it first suggested, may find for it another equally interesting and worthy end. For instance, an unruly ox (an eolith by this analogy) is readapted to a succulent use in the stew pot, and the tree that fails to bear fruit becomes firewood, both by sound eolithic principles of husbandry. The contrast is further sharpened in that eolithic craftsmanship is unique for each case, and therefore has little use for the uniform procedures and theories of design that the designing craftsman is committed to. Errors in the calculations of the designer may lead to the

Carl Hammer

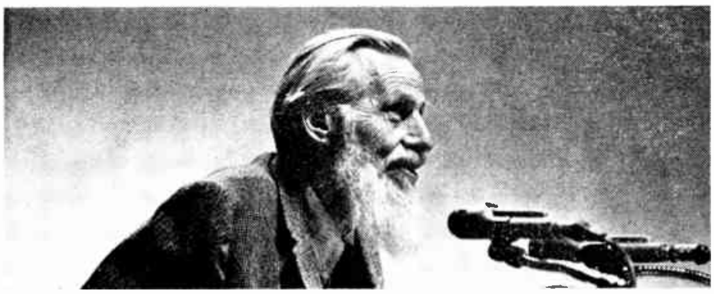
Yehoshua Bar-Hillel

Margaret Mead

collapse of a building. An eolithic builder, however, picks up his doorpost from the beach; and because it is many times more rugged than it needs to be, and the structure he erects so redundant in bracing, he need not worry about problems of stability in advance.

Why should we be concerned with eolithism? In one respect, the answer is clear. Our society has for many years been dominated by eolithism's opposite that all but excludes it, namely, by the principle of design.

As Hawkins points out, the logic of the design principle has steadily eaten into the organization of life as designers have gone from uniformity of materials to uniformity of parts and kits whose mass production eliminates all competing intermediate species. Worse, in the interest of mass production, men's goals—the purposes they set for themselves—are forced into restricted and stereotyped patterns, thus early exhausting the vital energies of individuals who attempt to overcome such restrictions. Up until recently, this tradition and this organization of our society has not been kind to those individuals desperate to get out of the specific jobs for which they have been trained.



Are things changing now? Is the pendulum beginning to swing back? Certainly, the restoration of eolithic principles, or something of this sort, in our social enterprises, and therefore in the lives of individuals, seems needed to redress the overemphasis on design based on uniformity; and indeed, there does seem to be evidence accumulating that favors such a trend. On the negative side, the exorbitant pollutions of all kinds caused by our machines are now evident to everybody who lives in the technological society and have aroused a universal though still helpless revulsion. On the positive side, the increasing cross-fertilization of professional and scientific disciplines offers new opportunities; the increasing interest of serious artists in working with engineers may bring much novelty and new purposes to technology as it has been conceived up till now. We might also regard the desire to devise computers so that individuals may play with them on-line in a trial-and-error fashion, to seek better ways of formulating problems, and to see problems in many aspects and on many levels, as a form of intellectual eolithism. The cybernetic movement itself, as it reaffirms its strength through the founding of a new society, signals a change.

We have chosen to stress here the Hawkins-Storm images of eolithism, although they are but the prelude to the weightier philosophic ideas. For Hawkins goes on to consider how individuals and society conceive their purposes and their futures. His examples raise the basic question as to the degree to which the state of a human being (or a society) can be conceived independently of its history or to which ends can be conceived independently of means. Without following his arguments, we must state simply that Hawkins comes to the point that we alone in the animal kingdom are able significantly to *set* purposes and not just seek them. Moreover, our purposes get destabilized and in the end reconstructed. He says that we are more purposive than other animals because we are able to analyze and resynthesize our goals into more comprehensive patterns of a hierarchical nature. On the other hand, we are less purposive than the other animals, for we are very conscious of the tentativeness and potential inadequacies of these designs. Hawkins sees the setting of goals as by nature an eolithic process. And the structure that an eolithic crafstman erects, as his response to the suggestions of the heterogeneous and varied nature of the materials, may be the life of the entire society. On the eve of each daring innovation, the eolithic craftsman may ask, "Am I quite sure I want to go that way?" And hard on this, Hawkins asks, "If our ancestors who invented the wheel or the smelter had foreseen us, would they not have shied away perhaps from the enterprise or at least thought twice about it?" That is, out of our heterogeneous resources, not all fit equally well to a given rationalized pattern of goals, and therefore some resources go unused or underdeveloped. He asks, "Where does the novelty come from that now and then intrudes itself into a standard situation and in the end transforms it into the beginning of a new pattern of work or conduct?" The rule that Hawkins offers is this: When in a dilemma, introduce novelty.

Which brings us to the point we mentioned at the outset, of conceiving choice as a constructive process. The whole situation in which we find ourselves, Hawkins concludes, in which means are so often the center of our interests and in which the ends, our purposes, are so fugitive as objects of serious discussion, might begin to be transformed provided we think of choice as a constructive process. "Then," says Hawkins, "we will perhaps be on the way to making the transition which I, for one, think we are deeply committed to in such a field as cybernetics. We are not only concerned with something called science, but also concerned with the setting of goals and, in particular, the setting of the goals that we need to set here and now."

It seemed generally agreed that the symposium had done what it intended—to set some intellectual fires at the dark edge.

For further information about the American Society for Cybernetics and publication date of symposium proceedings, write to: P.O. Box 9702, San Diego, Calif. 92109.

Photographs by George Tames.

REFERENCES

1. Wiener, N., *Cybernetics*, 2nd ed. New York: Wiley, 1961.
2. "Purposive systems: the edge of knowledge," *Proc. First Annual Symposium of Am. Soc. Cybernetics*, Oct. 26–27, 1967, Washington, D.C., to be published.
3. "Revolt against physics," editorial, *The New York Times*, p. 38, Feb. 12, 1968.
4. Bazelon, D. T., "The new factor in American society," *Proc. Conf. The Next Fifty Years 1967/2017*, Oct. 1–6, 1967, Am. Inst. Planners, to be published by Indiana University Press.
5. Amosov, N. M., *Modeling of Thinking and the Mind*, trans. by L. Finegold and L. J. Fogel. New York: Spartan Books, 1967, p. 162.
6. Licklider, J. C. R., "Man-computer symbiosis," *IRE Trans. Human Factors in Electronics*, vol. HFE-1, Mar. 1960, pp. 4–11.
7. Storm, H. O., "Eolithism and design," *Colorado Quarterly*, vol. 1, 1952.

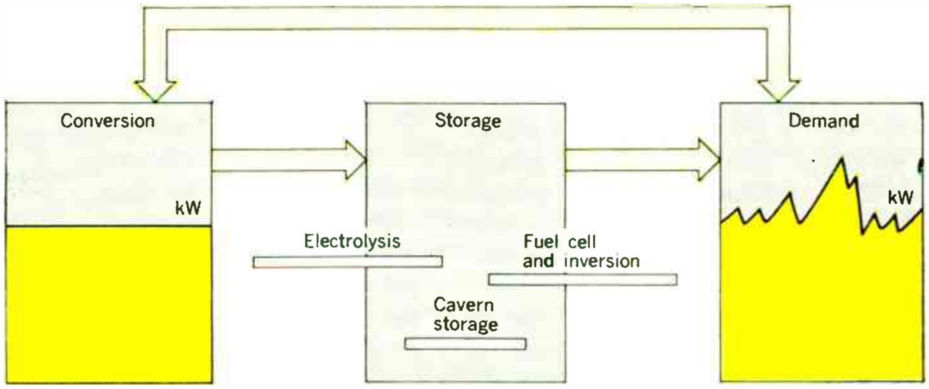


FIGURE 1. Model of three-macroblock system.

Economic optimization of energy conversion with storage

One important advantage in the combination of energy conversion and storage facilities is the realization of economies, on an annual basis, both in fuel costs and in required investment in plant

Arthur Bruckner, II *Louisiana State University*
W. J. Fabrycky *Virginia Polytechnic Institute*
James E. Shamblin *Oklahoma State University*

An experimental system being developed at Oklahoma State University represents the first phase of a research effort to formulate a model for optimizing an energy conversion and storage system. Some of the problems in providing energy storage as an alternative to added conversion capacity to meet peak demand are treated here. It is assumed that the conversion system is a steam or hydro plant with a controlled input as opposed to a solar or wind plant with a random input, although extension to the latter is possible.

Within the limits of present technology, fossil fuels, nuclear fuels, and hydro energy are converted through a series of processes to useful electric power, which is transmitted to and used by the consumer in essentially the same instant of time. Storage of energy is provided by nature in hydrocarbons, nuclear fuel, or in water behind a dam, and only the final conversion is performed by mankind. The fundamental problem in making use of energy sources other than fossil fuels, nuclear fuels, and hydroelectric energy results from the inability of man to gather energy at exactly the rate and quantity that he desires to use it. Thus, man must develop techniques to make his input and output energy rates independent of each other if he is to utilize sources of energy other than those presently stored by nature.

An experimental energy system with storage

Of the energy storage methods considered, it has been found that the electrolysis of water with subsequent storage of the gases is one of the most feasible. The system utilizing such a storage component is shown schematically in Fig. 1. The availability of an energy storage block allows utilization of energy sources other than those in which nature has stored converted solar energy. This technique does not, however, exclude those sources of energy in which nature has provided the initial conversion and storage. On the contrary, it will be shown here that energy storage capability can make the utilization of energy by these systems more economical.

In the experimental energy storage system,[1,2] energy sources are called upon to produce electric power by the use of solar cells, wind-driven generators, hydro turbines, or conventional steam power plants. The resulting current is used to electrolyze water, which, because of the closed nature of the process, produces hydrogen and oxygen gas at high pressure. The hydrogen may then be stored in steel tanks or in underground caverns. The oxygen may be similarly treated or used as a by-product.

An automatic control system is used to accelerate or decelerate the electrolysis process, based on the pressure in the hydrogen storage tank. The hydrogen in the pressure tanks is released through a pressure reducer to a fuel

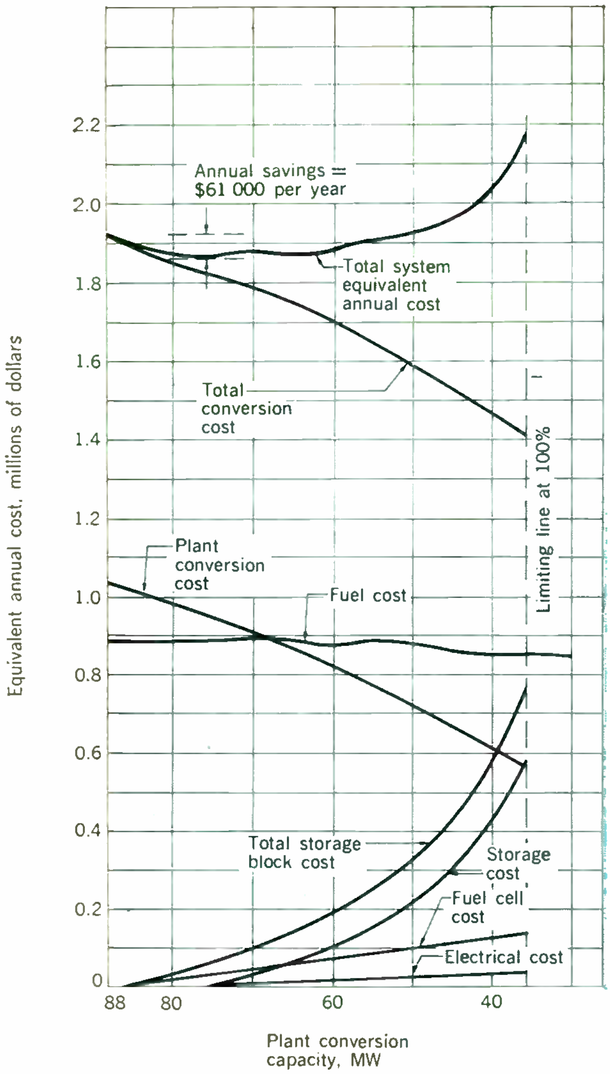


FIGURE 2. Optimization trade-off model.

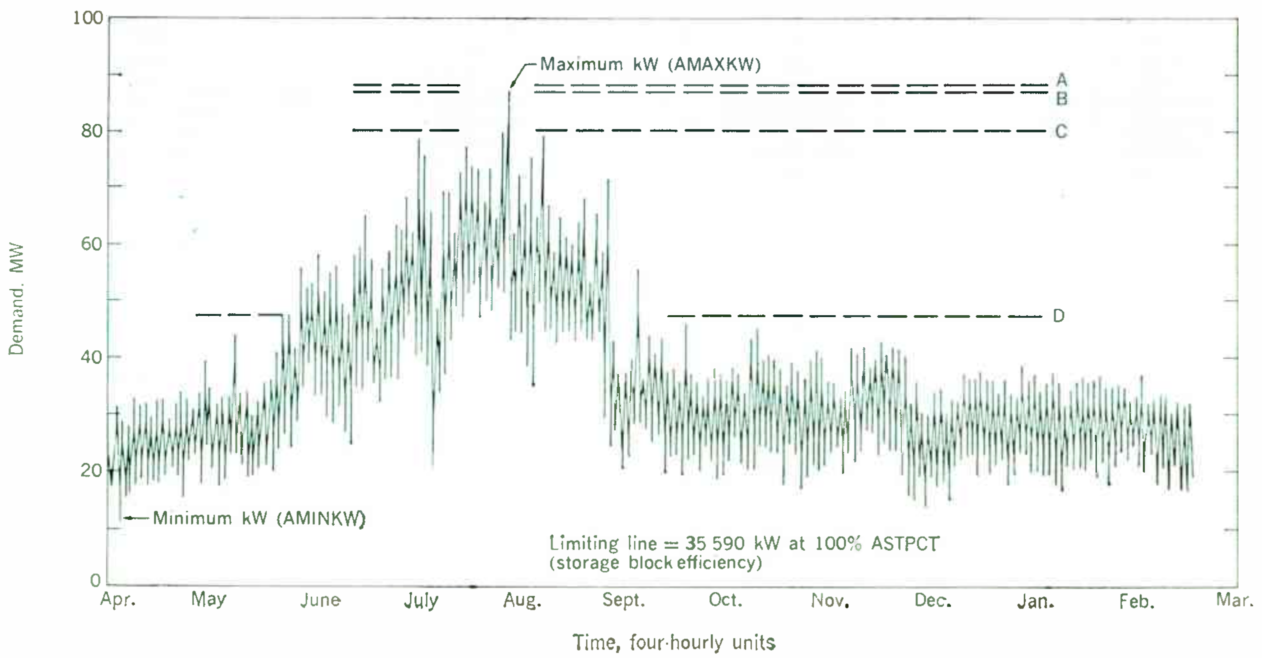
cell system. The amount of hydrogen released to the fuel cell is controlled by the load demand. Filtered air or pure oxygen must also be pumped into the fuel system, and the resulting water by-product must, of course, be recirculated into the pressure electrolysis system. Assuming that it is desirable to have reasonably high-voltage alternating currents, the fuel cell system output can be fed to an inverter and be made available to meet demand.

There is a distinct advantage of a system such as the one just described and illustrated in Fig. 1. During periods of low demand, excess power from the conversion block is used to store energy in the storage block. In such a system, the storage block supplies demand loads greater than the capacity of the conversion block. Thus, an energy conversion-storage system provides a means for making the demand for energy independent of energy supply; it is analogous to an inventory process.[3]

The optimization model

Briefly, Model I, the optimization model discussed in this article, may be described as follows: Total system equivalent annual cost is expressed as a function of conversion capacity as the controllable decision (design) variable. Two major cost components are involved. The first reflects the decreasing equivalent annual cost of energy conversion (steam plant, solar, hydro, nuclear, etc.) as the required conversion capacity or "scale of plant" decreases. The second reflects the increasing equivalent annual cost of energy storage capability needed to compensate for the reduction in conversion capacity for a specific storage efficiency. These major cost components, together with their subordinate cost components, are depicted in Fig. 2, which also illustrates the total system cost of a conversion–storage installation as a function of the energy conversion capacity. The ordinate intercept point gives the total cost of the conversion–storage system if no storage is provided and is, therefore, the cost of a conversion plant without energy storage. If a minimum point exists on the total cost

FIGURE 3. Demand function for electric energy.



function, a system incorporating energy storage is economically feasible. The conversion capacity at the minimum point is the least-cost "scale of plant." This is a design value useful in developing an optimal conversion–storage system, since this minimum point also establishes the energy storage component capacities required for the system. If no minimum point exists on the total cost function, it may be concluded that energy storage is not feasible for the specific application under study.

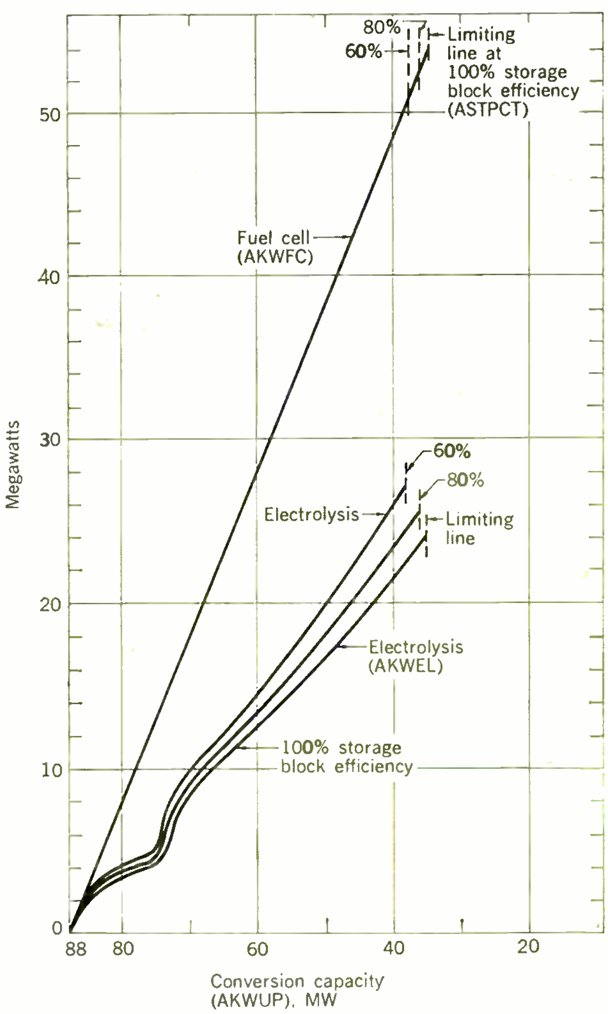
In the sections that follow, a representative situation will be presented to illustrate the foregoing model. Much of the computational work was performed on a digital computer. The output results are shown in the Appendixes. Specific values from the output are utilized in the analysis presented in the following sections.

The demand function

Demand for electric energy is the primary justification for the design and construction of an energy conversion and storage system. Because of the compulsion to meet demand, conventional conversion systems must provide a plant with sufficient capacity to meet peak demand. As a result, a corresponding high capital investment with low utilization occurs. Operating efficiencies are sacrificed when demand is less than plant capacity. With energy storage it is possible to make demand independent of supply, and thus economies result.

The demand function for the hypothetical example of this article is shown in Fig. 3. It spans a one-year period and exhibits a peak kilowatt demand that is approximately four times the minimum peak kW demand (when including daily variation, the peak-to-low demand ratio is approximately 8 to 1). This ratio is typical of a demand environment incorporating variations from daily, weekly, and seasonal consumer and industrial energy requirements. This particular function consists of 2190 points; one point (t) for every four hours. If demand data were plotted less frequently than every four hours, an error of significant magnitude would be introduced because of the lack of resolution between true kilowatt-hours and peak capacity. One-hour intervals would be more desirable. These 2190 points provide the basic data for the determination of system physical characteristics and the subsequent cost analysis. Conclusions concerning design of the system are based on this unique demand function with its individual capacity peak, total kWh, and particular pattern of demand.

FIGURE 4. Electrolysis capacity and fuel cell capacity vs. plant conversion capacity.



Evaluation of the storage block

Consider the line designated A in Fig. 3. This line represents the required capacity of the conversion block if no storage is provided. For the demand function of Fig. 3, this is 88 000 kW. Minimum kW capacity is 11 200 kW. Lines B, C, . . . represent decremental reductions in the capacity of the conversion block, which necessitate increased capacity in the storage block.

The ordinate of Fig. 3, from kW minimum (AMINKW) to kW maximum (AMAXKW), is divided into 100 such decrements. The results shown in Appendix A provide the basis for developing physical system characteristics. Column 1, ASTPCT, gives the efficiency of the storage block. Each value given in column 2, AKWUP, is the difference between kW maximum and successive kW capacity decrements within a given efficiency. These values are the decremental reduction in conversion capacity. For each conversion capacity considered, a given area (AONE) in kWh must be met by stored energy. This requires that certain kWh be stored (ATWO). Note that AONE = ATWO when the storage efficiency is 100 percent. The computations for a given efficiency terminate at a "limiting line," defined as that minimum conversion capacity below which it would not be possible to satisfy the demand requirements in kWh for the year.

The tabulated values in Appendixes A and B are given as a function of conversion capacity for a specific efficiency. Curves portraying these data may be developed. Their range for a given storage block efficiency indicates the effect of the appropriate limiting line.

Figure 4 exhibits the electrolysis capacity in kilowatts (AKWEL) as a function of the conversion capacity for various efficiencies of storage. It is developed from Appendix A by reference to columns 1, 2, and 8. Figure 4 also shows fuel cell capacity in kilowatts (AKWFC) as a function of plant conversion capacity. It is developed from Appendix A by reference to columns 1, 2, and 7. Finally, Fig. 5 shows the kilowatthours of energy (ATWO) needed to be stored as a function of conversion capacity. Information from Figs. 4 and 5 may now be used with information of Appendix B to develop cost extensions.

Cost extension computations

As portrayed in Fig. 2, conversion cost and storage cost make up a total cost function that may be used in decision making. This section is concerned with the economic analysis required in the development of these cost functions. It is based on the physical characteristics of the previous section together with assumed cost relationships.

These curves (Figs. 8, 9, and 10) are only hypothetical in order to demonstrate the capabilities and applications of this system model. The present level of the state of the art precludes cost information to the desired degree. However, the use of the curves does indicate the correct method for evaluating economic results. Moreover, the indicated approach enables the economic evaluation of cost/design parameters for relative merit in the total three-block system.

Conversion cost. Suppose that the equivalent annual installed cost of the conversion block is a function of its capacity in kilowatts, as given in Fig. 6. Note that the total cost increases at a decreasing rate. Although the function shown is not applicable to all conversion plants, it represents a typical steam or hydro plant. It is assumed here that the time value of money was considered.

The second major component of conversion cost is the cost of fuel if a steam plant is being considered. (A more sensitive fuel cost algorithm is now under study for Model II.) Appendix B gives the annual fuel cost as a function of the capacity of the conversion block and the efficiency of the storage block. The results are based on previously computed data of columns 2, 6, 9, and 10 in Appendix A, as well as the efficiency function of Fig. 7. When no storage is provided, the fuel cost is computed as follows: The demand function is divided into vertical segments; the average ordinate value of each segment determines the integer number of generators (four-generator plant) and the overall running load percentage. As an example,

A segment average ordinate intercept = 80 000 kW
Required integer number of generators = 4
Generators at 100% running load = 4 − 1 = 3
Kilowatts per generator = 22 000
Plant overall running load

$$= \frac{(80\,000 - 66\,000) + 3(22\,000)}{1(22\,000) + 3(22\,000)} = 91\%$$

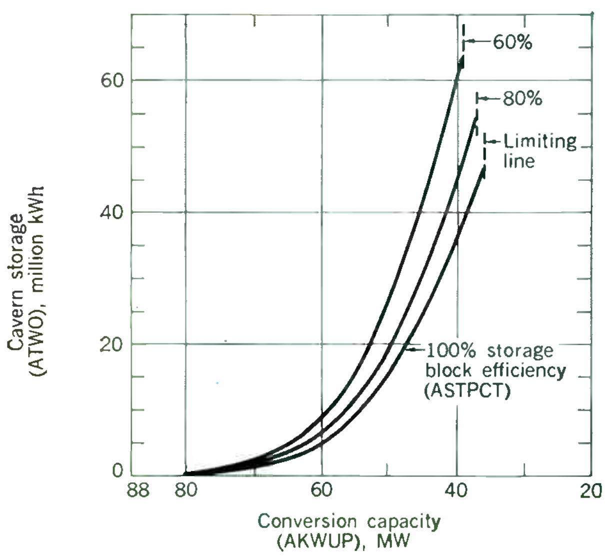


FIGURE 5. Cavern storage requirements as a function of plant conversion capacity.

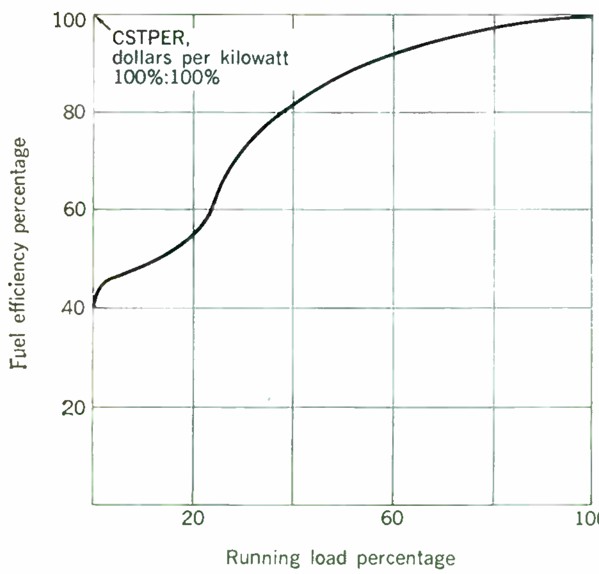


FIGURE 7. Fuel efficiency as a function of running load.

FIGURE 6. Power plant cost as a function of plant conversion capacity.

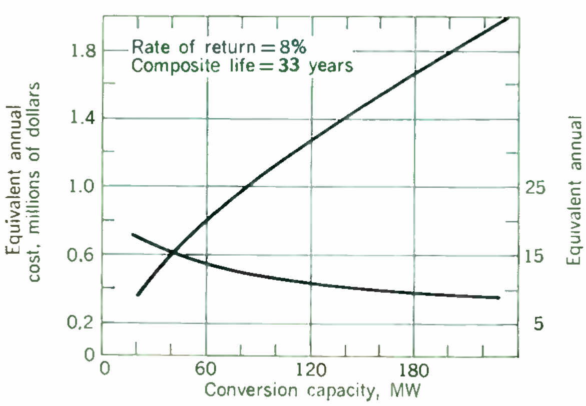
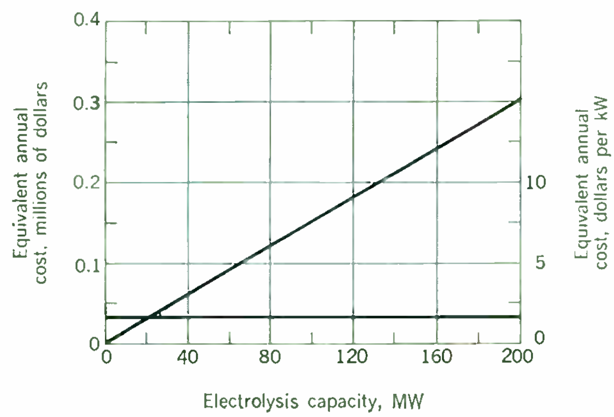


FIGURE 8. Electrolysis cost vs. electrolysis capacity.



By computer-lookup reference to Fig. 7, the overall plant fuel efficiency for the given segment is 99 percent. For this segment, the fuel cost is

$$\frac{\$ \text{ per kWh at } 100\%}{\text{efficiency } (\%)} \times \text{avg. kW} \times \frac{t}{\text{segment}} \times \frac{\text{hours}}{t}$$

When storage is used, a comparably equivalent adjustment is made for the kW per generator. The mix of generator kW is established as being equal respectively to the two plateaus of uniform running load (AKWUP, AKWDWN) until the limiting line of constant running load for the full year is reached. This represents the best design of plant in that uniform running load portions of the year are run at highest fuel efficiency. Load in kW per generator is established for each plateau such that the generator kW selected is most equivalent to the defined four equisized generators. The low plateau number of generators and kW per generator are determined by comparing AKWDWN per uniform kW per generator such that if the decimal portion of the ratio is greater than 0.5, the next highest integer number of generators is chosen. For this plateau, the kW per generator is then AKWDWN per number of generators. A similar method determines the balance of the generator mix and kW for AKWUP, except that the balance of required generation capacity is AKWUP − AKWDWN.

These segment costs are cumulatively added across that time portion of the conversion function where demand is variable. The total fuel cost for the year requires only the additional cost computation for the two plateau portions of the year that run at 100 percent efficiency. These expressions are as follows:

IDTUP = number of time periods of upper plateau

$$\text{Fuel cost up} = \text{AKWUP} \times \frac{\text{CSTPER}}{100\%} \times \text{IDTUP} \times \text{ADT}$$

IDTDWN = number of time periods of lower plateau

Fuel cost down

$$= \text{AKWDWN} \times \frac{\text{CSTPER}}{100\%} \times \text{IDTDWN} \times \text{ADT}$$

Therefore,

$$\text{Total fuel cost} = \text{cost up} + \text{cost down} + \Sigma \text{ segment fuel costs}$$

where

ADT = hours per time increment
IDTUP = number of upper plateau time increments
IDTDWN = number of lower plateau time increments
CSTPER = $ per kWh only at 100% efficiency (equals $0.0027/kWh in this study)

Note that there is an inverse relationship between storage efficiency and overall fuel efficiency. The latter increases as storage block efficiency decreases since the portion of the year at uniform running load is greater.

Storage cost. As indicated earlier, the storage block is composed of the electrolysis equipment, cavern storage, and the fuel-cell–inversion equipment. The cost of these components is assumed to be as shown in Figs. 8–10.[4]

Total system cost. Column 5 of Table I represents the sum of the component costs given in the previous columns; it is the total equivalent annual cost of the three-block system of Fig. 1 as a function of the conversion capacity of the system. Note that the total equivalent annual cost is $1 922 000 for no storage and that the conversion capacity required is 88 000 kW. As the conversion capacity decreases, the total equivalent annual cost decreases to a minimum value of $1 861 000. Beyond this point, the total equivalent annual cost begins increasing beyond bound. Thus, it may be concluded that the optimum conversion capacity for this situation is 74 940 kW if the efficiency of storage is 100 percent.

The entire system may now be designed from the parameters given at the minimum point. Specifically, it is required that the conversion block (plant) be 74 940 kW (with a generator mix level of 74 940 and 15 570 for the two plateaus in this case), that the capacity of the storage block be 52 224 kWh, that the capacity of the electrolysis equipment be 4377 kW, and that the capacity of the fuel-cell–inversion equipment be 13 050 kW. These design parameters, which are a direct result of the computations leading up to Table I, appear in Appendix A.

If the total equivalent annual cost function has no minimum point, a system without energy storage is the least-cost system. Such a condition may be forced by a storage block efficiency that is sufficiently low.

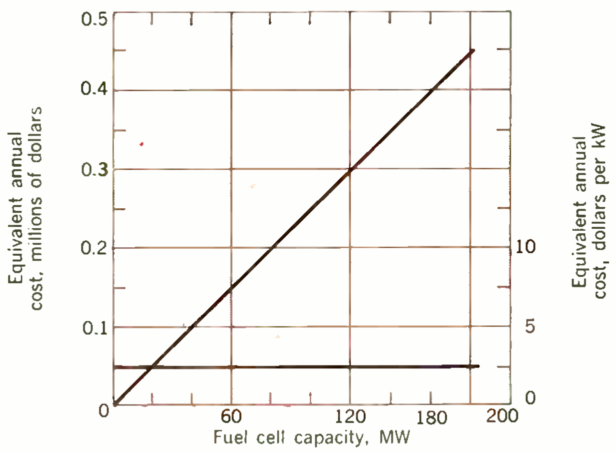
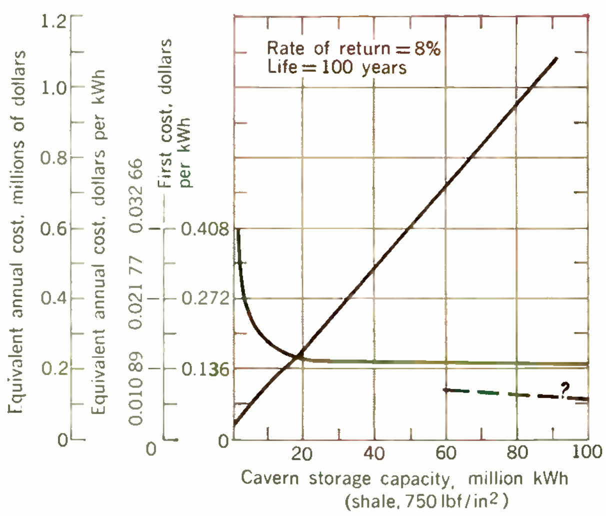


FIGURE 9. Fuel cell cost vs. fuel cell capacity.

FIGURE 10. Cavern storage cost vs. storage capacity.



I. Total system equivalent annual cost for a storage block efficiency of 100 percent (partial listing of computer tabulation cost extensions)

Plant Conversion Capacity (kW)	Annual Cost of Conversion Block (000's)	Annual Fuel Cost (000's)	Annual Cost of Total Storage Block (000's)	Total System Equivalent Annual Cost (000's)
88 000	$1040	$882	$ 0	$1922
87 230	1030	882	3	1915
84 160	1010	881	14	1905
81 080	994	875	24	1893
78 010	963	880	32	1875
74 940	944	875	42	1861
71 870	920	903	81	1904
68 800	890	894	112	1896
65 720	860	884	131	1875
62 650	835	877	161	1873
59 580	804	869	194	1867
56 510	780	908	232	1920
53 440	743	889	279	1911
50 360	719	872	329	1920
47 290	680	861	385	1926
44 220	660	853	460	1973
41 150	620	848	549	2017
38 080	585	847	650	2082
35 590	563	854	745	2162
Limiting line				

Design application summary

The trade-off optimization model presented provides a means for determining design parameters for an energy conversion–storage system. The supporting computations were performed on a small-scale digital computer. This generalized modeling, often called system simulation, is amenable to computation by a digital computer.

Perhaps the prime advantage of system simulation is its capabilities for predicting the effects of parameter changes on the total system, while at the same time using sufficient empirical data to describe an actual system. The Model I can analyze any actual demand curve with individual annual cost parameters for conversion capacity and other equipment. Fuel cost can be treated individually on a regional basis.

More important, however, is the capability of this model to assist decision making for both system physical requirements and economic values. For example,

1. What is the economic limit to use of storage versus the physical limit (limiting line) for a particular demand curve?
2. What is the effect on the minimum cost point of the total system by an increase in fuel cell efficiency?
3. If one component of the system can achieve unusually high performance by a high-price part, will the system gain warrant the marginal part cost?
4. How much fuel and fuel cost can be saved by a uniform running load for an individual demand curve?
5. What are the capacity specifications for an electrolysis unit in a plant design for a given demand curve and an expected efficiency level?
6. For which component or subcomponent will additional research or operating control efforts offer the best potential yield in savings?
7. What will be the effects on a plant design and cost of operation if the shape of the demand curve is affected by an industrial consumer?

Conclusion

The preceding sample of questions can be directly relevant to marginal-return studies of power plant design. These types of questions can be examined by utilization of this model. For example, though based on this study's given cost-parameter values, the indications are nevertheless that the cavern storage component of the storage block is significant. Hence, another area requiring research is the more economic storage of hydrogen whether in natural or man-made caverns.

Perhaps the most significant feature of this model's design is the evaluation of energy requirements on an annual demand basis. Typically, operating efficiency studies with some temporary storage are made on a more daily basis for fuel economies. However, these possible savings have little effect on size-of-plant costs except for the few days a year at peak load. Eventually, annual energy conversion with storage could realize economies in both fuel costs and plant investment.

Even though the present system state of the art is not yet sufficiently advanced for commercial application, energy storage systems do offer some potential advantages. Generally inherent is the ability of such systems to come up to load level rapidly and independently in case of power blackouts. Storage systems combined with nuclear plants could offer the joint advantages of a relatively constant running load and a much smaller-sized plant for the same demand requirements.

A comment on future research activities in energy conversion and storage is pertinent before concluding. First of all, the Model I design is the result merely of the initial phase of studies in this general area. It is limited to a steady-state determination of demand and is applicable only to conventional conversion systems (controlled input). Following adaptation of the program to a larger-scale computer, a number of modifications are immediately planned. One area concerns the logic of storing energy throughout the year. Alternative algorithms will be considered. The new model will include a greater system-block breakdown for examination of many individual storage components. On a broader basis, simulation model development is planned to permit the investigation of energy conversion with storage systems operating under random, or uncontrolled, input power sources.

Full text of a paper recommended by the Power Generation Committee of the IEEE Power Group for presentation at the IEEE Winter Power Meeting, New York, N.Y., Jan. 28–Feb. 2, 1968.

REFERENCES

1. "Energy storage: key to our economic future," Energy Storage Research Team, Office of Engineering Research, Oklahoma State University, Stillwater, 1963.
2. Thornton, R. T., and Hoke, T., "Storage as an economic alternative for conventional peaking capacity," *Proc. 2nd Ann. Conf. Energy Conversion and Storage*, Office of Engineering Research, Oklahoma State University, 1964.
3. Fabrycky, W. J., "Energy storage as an inventory process," *Proc. 1st Ann. Conf. Energy Conversion and Storage*, Office of Engineering Research, Oklahoma State University, 1963.
4. Norton, J. R., "Methods and economics of high-pressure gas storage," *Proc. 1st Ann. Conf. Energy Conversion and Storage*, Office of Engineering Research, Oklahoma State University, 1963.

1	2	3	4	5	6	7	8	9	10
ASTPCT	AKWUP	AOUE	[illegible]	A2THEO	AKWDWN	[illegible]	AKWEL	IDTUP	IDTDWN
1.0000	.8723E 05	.30720000E 04	.30720000E 04	.30720000E 04	.1196E 05	.7680E 03	.7680E 03	1	1
1.0000	.8646E 05	.61440000E 04	.61440000E 04	.61440000E 04	[illegible] 05	.1536E 04	.1536E 04	1	1
1.0000	.8569E 05	.92160000E 04	.92160000E 04	.89600000E 04	.1312E 05	.2304E 04	.1920E 04	1	2
1.0000	.8492E 05	.12288000E 05	.12288000E 05	.12032000E 05	.1350E 05	.3072E 04	.2304E 04	1	2
1.0000	.8416E 05	.15360000E 05	.15360000E 05	.15488000E 05	.1373E 05	.3850E 04	.2534E 04	1	5
1.0000	.8339E 05	.18432000E 05	.18432000E 05	.18560000E 05	.1388E 05	.4608E 04	.2688E 04	1	5
1.0000	.8262E 05	.21504000E 05	.21504000E 05	.21632000E 05	.1404E 05	.5376E 04	.2841E 04	1	5
1.0000	.8185E 05	.24576000E 05	.24576000E 05	.24704000E 05	.1419E 05	.6144E 04	.2995E 04	1	5
1.0000	.8108E 05	.27648000E 05	.27648000E 05	.27776000E 05	.1434E 05	.6912E 04	.3148E 04	1	5
1.0000	.8032E 05	.30720000E 05	.30720000E 05	.32896000E 05	.1450E 05	.7680E 04	.3302E 04	1	10
1.0000	.7955E 05	.33792000E 05	.33792000E 05	.39040000E 05	.1465E 05	.8448E 04	.3456E 04	1	10
1.0000	.7878E 05	.36864000E 05	.36864000E 05	.45542400E 05	.1480E 05	.9216E 04	.3609E 04	1	11
1.0000	.7801E 05	.39936000E 05	.39936000E 05	.52300800E 05	.1496E 05	.9984E 04	.3763E 04	1	11
1.0000	.7724E 05	.43008000E 05	.43008000E 05	.59059200E 05	.1511E 05	.1075E 05	.3916E 04	1	11
1.0000	.7648E 05	.46080000E 05	.46080000E 05	.67225600E 05	.1527E 05	.1152E 05	.4070E 04	1	16
1.0000	.7571E 05	.49152000E 05	.49152000E 05	.77248000E 05	.1542E 05	.1228E 05	.4224E 04	1	18
1.0000	.7494E 05	.52224000E 05	.52224000E 05	.88307200E 05	.1557E 05	.1305E 05	.4377E 04	1	18
1.0000	.7417E 05	.41690400E 05	.41690400E 05	.42082240E 06	.1795E 05	.1382E 05	.6758E 04	34	59
1.0000	.7340E 05	.52539200E 06	.52539200E 06	.53534400E 06	.1841E 05	.1459E 05	.7219E 04	39	80
1.0000	.7264E 05	.64648000E 06	.64648000E 06	.63615040E 06	.1872E 05	.1536E 05	.7526E 04	43	84
1.0000	.7187E 05	.78113600E 06	.78113600E 06	.76962560E 06	.1911E 05	.1612E 05	.7910E 04	48	91
1.0000	.7110E 05	.93236800E 06	.93236800E 06	.93782720E 06	.1949E 05	.1689E 05	.8294E 04	5[illegible]	122
1.0000	.7033E 05	.10942480E 07	.10942480E 07	.10912864E 07	.1980E 05	.1766E 05	.8601E 04	53	126
1.0000	.6956E 05	.12628240E 07	.12628240E 07	.12597408E 07	.2010E 05	.1843E 05	.8908E 04	58	154
1.0000	.6880E 05	.14480400E 07	.14480400E 07	.14577440E 07	.2041E 05	.1920E 05	.9216E 04	65	166
1.0000	.6803E 05	.16580640E 07	.16580640E 07	.16642944E 07	.2072E 05	.1996E 05	.9523E 04	72	173
1.0000	.6726E 05	.18962880E 07	.18962880E 07	.19052480E 07	.2103E 05	.2073E 05	.9830E 04	82	203
1.0000	.6649E 05	.21538240E 07	.21538240E 07	.21599456E 07	.2133E 05	.2150E 05	.1013E 05	84	214
1.0000	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	248
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1.0000	.3808E 05	.40493240E 08	.40637358E 08	.40637358E 08	.3424E 05	.4992E 05	.2304E 05	678	1302
1.0000	.3731E 05	.42638160E 08	.42638160E 08	.42661612E 08	.3462E 05	.5068E 05	.2342E 05	715	1331
1.0000	.3654E 05	.44900472E 08	.44900472E 08	.44719569E 08	.3500E 05	.5145E 05	.2380E 05	763	1348
1.0000	.3577E 05	.47308568E 08	.47308568E 08	.47234102E 08	.3546E 05	.5222E 05	.2426E 05	808	1377
1.0000	.3559E 05	.47906832E 08	.47906832E 08	.48083474E 08	.3562E 05	.5240E 05	.2442E 05	820	1392
.9000	.8723E 05	.30720000E 04	.34133333E 04	.33792000E 04	.1204E 05	.7680E 03	.8448E 03	1	1
.9000	.8646E 05	.61440000E 04	.68266666E 04	.71168000E 04	.1288E 05	.1536E 04	.1689E 04	1	2
.9000	.8569E 05	.92160000E 04	.10240000E 05	.10188800E 05	.1327E 05	.2304E 04	.2073E 04	1	2
.9000	.8492E 05	.12288000E 05	.13653333E 05	.13952000E 05	.1365E 05	.3072E 04	.2457E 04	1	5
.9000	.8416E 05	.15360000E 05	.17066666E 05	.17024000E 05	.1381E 05	.3840E 04	.2611E 04	1	5
.9000	.8339E 05	.18432000E 05	.20480000E 05	.20096000E 05	.1396E 05	.4608E 04	.2764E 04	1	5
.9000	.8262E 05	.21504000E 05	.23893333E 05	.23168000E 05	.1411E 05	.5376E 04	.2918E 04	1	5
.9000	.8185E 05	.24576000E 05	.27306666E 05	.27776000E 05	.1434E 05	.6144E 04	.3148E 04	1	5
.9000	.8108E 05	.27648000E 05	.30720000E 05	.32896000E 05	.1450E 05	.6912E 04	.3302E 04	1	10
.9000	.8032E 05	.30720000E 05	.34133333E 05	.39040000E 05	.1465E 05	.7680E 04	.3456E 04	1	10
.9000	.7955E 05	.33792000E 05	.37546666E 05	.45542400E 05	.1480E 05	.8448E 04	.3609E 04	1	11
.9000	.7878E 05	.36864000E 05	.40960000E 05	.52300800E 05	.1496E 05	.9216E 04	.3763E 04	1	11
.9000	.7801E 05	.39936000E 05	.44373333E 05	.59059200E 05	.1511E 05	.9984E 04	.3916E 04	1	11
.9000	.7724E 05	.43008000E 05	.47786666E 05	.67225600E 05	.1527E 05	.1075E 05	.4070E 04	1	16
.9000	.7648E 05	.46080000E 05	.51200000E 05	.77248000E 05	.1542E 05	.1152E 05	.4224E 04	1	18
.9000	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	18

Appendix A

Appendix B

WHEN NO STORAGE IS USED AT AMAXKW = .8800E 05 THE COST OF FUEL IS EQUAL TO .88188520E 06 WHICH IS INDEPENDENT OF ANY LEVEL OF STORAGE SUB-SYSTEM (BLOCK) EFFICIENCY.

ASTPCT	AKWUP	CSTFOF
1.0000	.8723E 05	.88216612E 06
1.0000	.8646E 05	.88322688E 06
1.0000	.8569E 05	.88314902E 06
1.0000	.8492E 05	.88199749E 06
1.0000	.8416E 05	.88093500E 06
1.0000	.8339E 05	.87900568E 06
1.0000	.8262E 05	.87827004E 06
1.0000	.8185E 05	.87723795E 06
1.0000	.8108E 05	.87491520E 06
1.0000	.8032E 05	.87435777E 06
1.0000	.7955E 05	.88253678E 06
1.0000	.7878E 05	.88124167E 06
1.0000	.7801E 05	.88008848E 06
1.0000	.7724E 05	.87874161E 06
1.0000	.7648E 05	.87776239E 06
1.0000	.7571E 05	.87567023E 06
1.0000	.7494E 05	.87505509E 06
1.0000	.7417E 05	.87341562E 06
1.0000	.7340E 05	.90710251E 06
1.0000	.7264E 05	.90586075E 06
1.0000	.7187E 05	.90288994E 06
1.0000	.7110E 05	.90023436E 06
[illegible]	[illegible]	[illegible]
1.0000	.3808E 05	.84737797E 06
1.0000	.3731E 05	.84718465E 06
1.0000	.3654E 05	.84649729E 06
1.0000	.3577E 05	.84659412E 06
1.0000	.3559E 05	.85417839E 06
.9000	.8723E 05	.88251150E 06
.9000	.8646F 05	.88355969E 06
.9000	.8569E 05	.88330409E 06
.9000	.8492E 05	.88194887E 06
.9000	.8416E 05	.88093915E 06
.9000	.8339E 05	.87935429E 06
.9000	.8262E 05	.87854323E 06
.9000	.8185E 05	.87724625E 06
.9000	.8108E 05	.87573086E 06
[illegible]	[illegible]	[illegible]

Out of the growing need to better man's grasp of the real world in today's technological revolution, game theory has emerged. It has given us a new tool upon which we can rely when decisions become difficult. The purpose of this article is to enlighten the reader about a problem with which he might never come in contact, but nevertheless one as important as he might ever encounter; namely, how to cope with the rapidly changing military warfare of our times—where thousands or millions of people's lives are in the balance. This discussion of gaming draws heavily upon the defense problem, and two methods are developed in defining solutions for the theoretical model.

Game-theoreti

Games of strategy are important!

Consider for a moment, if you will, some important situation for which you, as an adversary, do not know the outcome of your actions before you begin them. A moment's thought should reveal a great number of such situations. But let's not dwell on these; let's think of decisions on which the life or death of millions or thousands of people or even just one person rides on the correct action. This is part of the problem in gaming.

Note that, under this guise, a most important fact about game theory has pushed its way outward. This fact is called value or objective. In order to qualify as a game, the objective(s) of the adversaries must be established. Now then, certain aspects of the play are considered more important than others. These aspects are essentially what we are searching for in gaming. As can be easily seen, though, the importance of a play depends very directly on the rules by which the game is played.

Take chess, for example. Suppose the objective of each side is to win the game. The rules state that in order for one to win he must mate his opponent's king or force him to resign (that is, the opponent concedes that the mate is inevitable). One result of this rule is that "safety of the king" is an important principle. Were the rules changed to force one side, in order to win, to capture *all* the opponent's pieces, I dare say that the corresponding principle of safety of the king would be severely modified. Note also that a change in the objective forces changes in the importance of the plays. For example, you can play the game to have as a result a "good" game. The rules, unfortunately, are not given for this objective, hence the principles one can derive from this are limited. This does not imply that there are *no* rules, because in fact there are, but it would take an application of psychology or perhaps philosophy to discover them. It is interesting to note that these rules (laws of nature) that man is seeking are being mirrored by the importance with which he attributes his plays.

In the subject of real-life games, consider a war game, in which the question is what weapon should be used on a particular target. Military strategists who seek answers to this question should really ask, "Assuming that the target's destruction is important, what is the best means by which we can destroy it while taking into consideration the overall objective?"

Some of these considerations might include

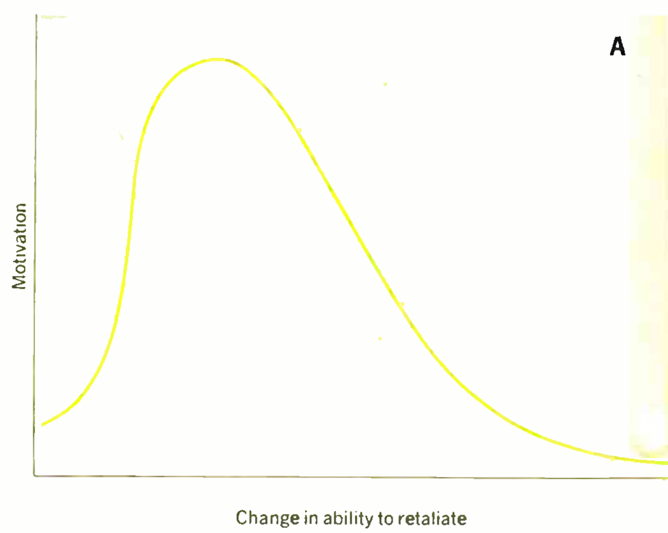
1. Cost of weapons able to destroy the target.
2. Side effects.
3. Counterattack possibilities.

In addition, because of uncertainty, some degree of "confidence" must be attached to the answer. Righteousness of means, although a factor, is not very measurable and, as such, is relegated to someone higher up. We must ignore it or code around it in computer models.

Let us carry the problem further. Cost is a factor that, at best, is difficult to describe, but with which strategists seem to be most at home. This is mainly due to the simplification that every play can be reduced to a dollars-and-cents value or, if necessary, to a number of people *not saved*, or to some other convenient measure. Further, since during destruction at least one stockpile is drained at every play (namely, the weapon), a minimization of this cost is usually employed in the investigation of alternatives.

On the other side of the coin, we have retaliation. This is where the gaming comes in. Two factors influence strategist's play: ability and motivation. In fact, since it is a reaction, each play can be expected to generate its own reaction based upon one's ability or motivation. Motivation, in some respects, is based upon the change in the ability to retaliate. Figures 1(A), 1(B), and 1(C) give a possible representation of this dependence. Figure 1(C) is a composite of the other two.

This last graph reduces motivation and ability to react to a cost equation. This procedure (or one like it) is always followed by strategists before they make their decisions. It is essential if they are going to make any comparison of the opposing forces. Note here that it is beginning to resemble a game with only offensive-type



applications

The age-old doctrines of warfare, while as realistic as ever, have been subjected to the rigorous disciplines of gaming theory, psychology, and missile strategy in an effort to maintain a persuasive balance of power

James P. Dix *Control Data Corporation*

weapons; each side alternates moves until one wins. Add to this such complications as defensive weapons, passive shelters, backfiring, and immoral acts (killing civilians, etc.) and you have some of the elements of a brilliant (and important) game.

Thus, in the back–back lines (in the United States), a game of strategy and tactics is taking place long before the play is made or, for that matter, before *any* play is made.

Consider the decision that Secretary MacNamara had been facing for about six years—whether or not to deploy an antimissile defense system and, if so, what size. At first, the decision was negative because of our lack of capability. Indeed, the Nike Zeus system could not successfully defend against the decoy-type attack. With the advent of Nike X and its extremely costly budget, it was necessary to weigh cost with the need to defend people. The Secretary hedged on the decision to deploy that weapon system by saying that it was not the time to decide. Then, under the influence of President Johnson, he explained that there is a strong chance that Russia (considered as the main opponent) and the United States could be locked in an arms race which would not only lead to great worldwide instability, but would all be wasted effort because of the possibility of total disaster. Our side is even going so far as to throw in a tactical shot by convincing the other side of the futility of such efforts on either side. This is true gamesmanship!

We have come to an era in warfare in which decisions are based, not on the outcome of the next battle or the next campaign, but on the entire complex of the socio-political situation years in advance. The decision to "escalate" the war in Vietnam is a difficult one. There are many and subtle nuances to each decision. The blockading of the Haiphong port, for instance, could spell out disaster if Russia were motivated to retaliate.

Decisions such as these must utilize tools that heretofore have never been used. In the vastness of the technological expansion, many questions become difficult to answer. Some typical ones are: What are the plays in this game? How do we know if we are capable of making any play? How do we know that, if we make the play, it will succeed? To date, we have no real experience in these matters; hence, we must trust our delegated technical advisors.

The solutions to these problems are, in part, the subject of this article. The methods of solutions discussed are, by no means, the only ones available; however, these are the ones with which I am most familiar. My experience dates back to a tour of service at Colorado Springs, Colo. There, at the headquarters of the Army Air Defense Command, one of the major tasks was to deploy a Nike X weapons system at all major cities around the continental United States. Depending on the amount of dollars available, we were able to deploy X number of sites. I cannot go into great detail about the system, since much of it is still classified, but I will discuss methods that were known in solving this problem.

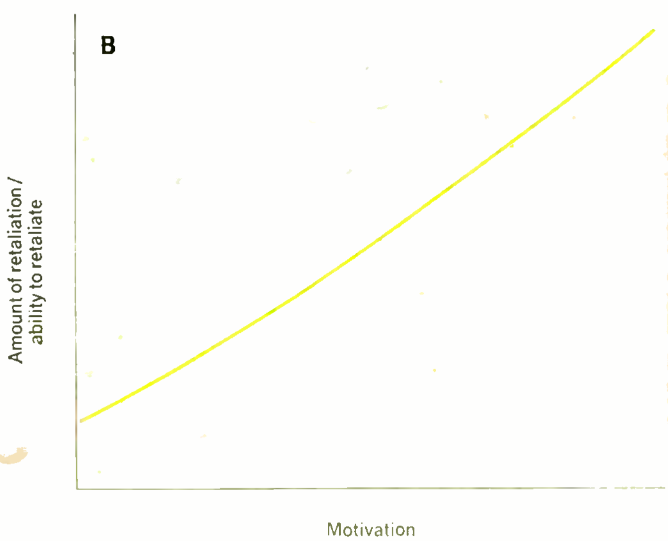
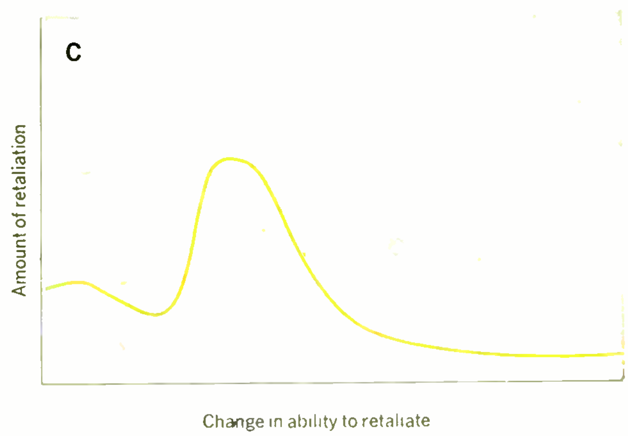


FIGURE 1. Various displays exhibiting dependence of motivation and ability to retaliate in a defensive strategy.



The game-theoretic model

The game-theoretic technique proposes an exact solution—the so-called minimax. In most cases, however, the problems are so complex that we cannot delimit all the strategies; hence we are reduced to merely finding the important principles.

Since I will be using some terms with which many readers are not familiar, some brief definitions follow in order that they can be used as a reference:

Firing doctrine. A set of rules by which the defense assigns missiles to objects and determines where and when to launch them.

Cloud pattern. As an ICBM (or other ballistic missile) is launched from a site that is far away, it may break up in a predetermined fashion as it approaches the target. The result is called a "cloud pattern" or a "cloud" and it contains warheads, decoys, scrap metal, jamming devices, and other tactical devices. The purpose of the defense, therefore, is to single out the warheads from the rest.

Fusing distribution. A term used by the offense when setting up the timing for the actual fusing and detonation of the warhead. It is usually rigged by altitude; hence it can be called a fuse altitude distribution.

Payoff per warhead. The extent of damage a warhead achieves. Over many warheads, this represents a goal for both sides. The offense wants as high a payoff per warhead as it can obtain, while the defense trys to limit the payoff to the smallest value *it* can obtain.

Price of admission. The number of ICBMs necessary to accomplish a specific goal.

Penetration. The ability of the offensive warhead to burst at an altitude low enough to do more than a specific amount of damage.

In examining the alternatives for selecting a certain firing doctrine, one must consider that the opposition knows exactly what cloud pattern and fusing distribution are best against each one. Now, knowing this, the firing doctrine is selected so that, no matter which cloud pattern and fusing distribution are chosen, the enemy will be able to do no better than the worst of the bests. Let us consider some of the important aspects of both sides.

First of all, the objective of the defense is to defend as many people as possible. This objective could be changed to defend as many goods or retaliation capabilities or whatever. The point is that the assumption influencing the offense's objective must be such that he desires to destroy as many people or as many goods or as many retaliation capabilities [or whatever the defense (our side) has as an objective] as possible. If not, each is essentially fighting a separate war, in which case one could win or lose depending on whose viewpoint is taken. Although this might happen in reality, it is clear that if one is not prepared for all of the most probable eventualities, the opposition can have a crushing effect. The mathematical equation describing this situation is:

$$v(o) = -v(d) \quad (1)$$

where $v(o)$ = measure of offensive value and $v(d)$ = measure of defensive value. This is often called a zero-sum game, since $v(o) + v(d) = 0$.

The second assumption to be made is that, with a fixed number of dollars for the defense and an infinite number of dollars for the offense, the offense must inevitably overcome the defense. Stating this, I must add that the use of fixed defense dollars is the only practical way of judging the merits of the system because, in this case, once they are deployed, the sites are fixed and the opposition can then observe and adjust before making its attack. Various fixed-dollar defenses can be tried to determine which is the most attractive. Moreover, it is not practical to vary this dollar amount in any convenient manner because of the method by which appropriations are made in Congress. This rule restricts play a great deal, but in turn makes the decisions easier. This second assumption, then, implies that you want to get the most for your money. That is, your objective is to find the deployment that makes the offense pay the highest price of admission at a reasonable cost to your side. This approach is often referred to as optimizing the cost/effectiveness ratio.

A third assumption is that, due to integral restrictions, some cities will not be defended at all (still assuming that our objective is to defend people). In other words, since you cannot spread a local defense system so thin that it protects everyone (there is no such thing as half an ABM), there will be one city that is ranked highest (in population) among those that are undefended. This city is usually used as the "key." The principle is to avoid *overprotecting* the large cities, since the offense can get what they want (that is, what the defense doesn't want them to have) at minimum cost by going to the undefended cities. Similarly, don't underprotect the large cities. A balanced defense then is the objective. By noting the key city's population, one can then say that the objective of a balanced defense is to limit the offense to a payoff per offensive warhead less than or equal to this population.

$$P_w \leq POP(k) \quad (2)$$

where k is the number of the first undefended city. This is a restriction that, when coupled with the next, states that you must achieve a sufficiently strong defense to make the enemy's price of admission $P_a \geq POP(i)P_w$, where $POP(i)$ is the population of the ith local defense (if any). This assumption states that once the offense has "penetrated," he has wiped you out (that is, one good blow is sufficient to incapacitate the defense). It follows that the missile defense system must be extremely reliable and accurate as long as there is a stockpile. It further states that penetration depends entirely on defensive stockpiles of ABMs rather than on any inherent weakness. This assumption is, in part, based on the fact in nuclear warfare that, in general, surface damage increases the deeper the penetration.

With these basic assumptions, we then draw up a plan for setting up a game-theoretic model that will lead us to solve the problem of siting. Because of the previously stated correlation regarding penetration and stockpile, it turns out that the best plan is to give each city a fixed number of ABMs to deploy as it chooses. Further, since sites are governed by terrain and government-owned or acquirable land, the number of plays is fairly limited.

The next question to ask is: Given a deployment, how good is it? That is, given a sufficient stockpile for the assigned site locations, could it adequately defend the city whose population is distributed in some known manner? In other words, where are the intercepts (if any) going to take place and who is going to do the intercepting?

This question is a natural one since (a) the intercept has to take place, and (b) because of the objective of a balanced defense, it cannot allow a penetration below the P_w line. The offense wants to know the answer to that

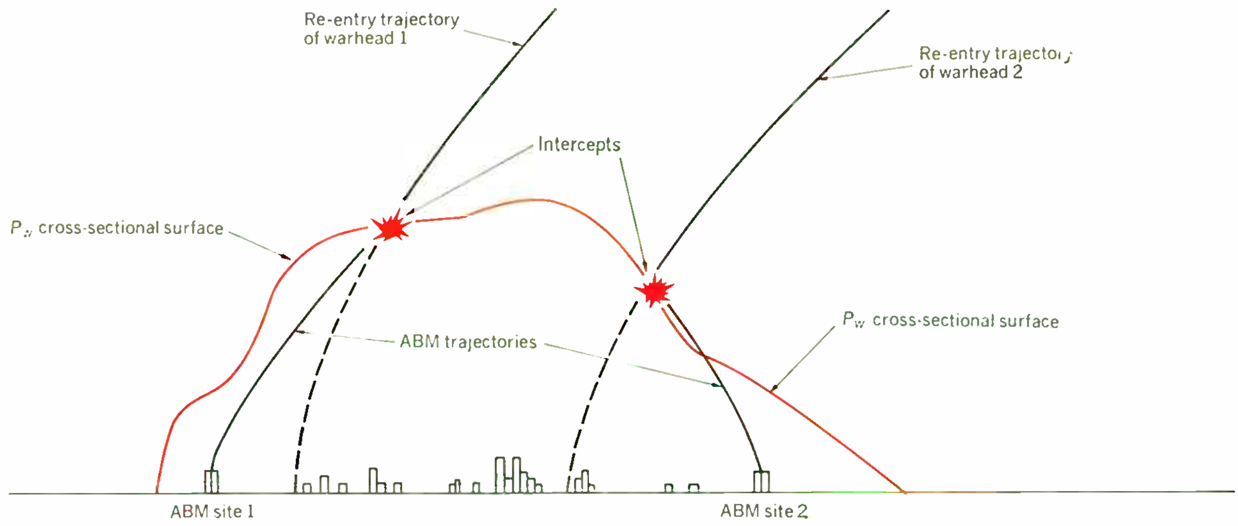


FIGURE 2. Fixed-site defensive strategy for the game-theoretic model.

question, too. He is anxious to get his burst off, but would also like to penetrate the P_w surface. This is that surface which, if a burst of the assumed weapon took place, would just kill P_w people. Since, in order to make an intercept, you must first "see" your target, you must wait as long as possible before making the shot. Also, because accuracy increases as the intercept becomes closer, you must wait longer. But this introduces a dilemma, since (a) and (b) recommend getting him as soon as possible.

Figure 2 illustrates the defense's plan. It turns out, after some rigorous mathematics using game theory, that (under all the assumptions) if you select your intercepts exactly on the P_w surface, you can best determine which sighting scheme prevails.

In this simplified model, this last statement is surely a significant one since with it one may obtain an extremely valuable siting criterion. In other words, sites are placed so as to be nearest the intercepts. Further, because the intercept (P_w) surface is fixed, insofar as the size of a warhead is fixed, you have almost solved your problem.

There are overriding criteria, however, all of which push this surface higher and farther away. Since these are tactical rather than strategical, they cannot be given too much weight and are not given much coverage in this article. For those who are interested, they include firing doctrine requirements like avoiding fratricide (killing one's own) and bad cross angles (angle of attack versus ABM angle of intercept), offensive re-entry maneuvering vehicles, and radar blackouts due to precursor bursts.

Before delving into various computer models, a few words about offensive threats must be stated.

One must defensively be capable of withstanding all reasonable attacks. Any weakness will surely be discovered and all efforts to defend a city will be wasted.

What are some of the options of attack that the aggressor has?

1. He can vary his warhead size or the actual number of warheads. The deterring factor, of course, is the increased weight of the ICBM (or other type of ballistic missile). Weight and cost factors are given in Figs. 3(A) and 3(B). The payoff might be worth the risk, however; see Figs. 3(C) and 3(D).

2. He might vary his re-entry angle. Here, too, his penalty is weight, as demonstrated in Fig. 4(A). The payoff curve is shown in Fig. 4(B).

3. He might vary his re-entry azimuth. The weight penalty is severe in this case; that is, you either fire the great circle or don't fire (see Fig. 5).

4. Another possibility is to select various aim points in a city to receive one's damage (Fig. 6).

Attempting to put all these considerations together is a formidable task, but not an unsolvable one (at least to the satisfaction of most). I plan to resolve the problem in two ways.

The first is so simple that it might mislead by its very simplicity. It tackles the problem by finding one or two key factors that can be computed moderately easily. These factors can be arrived at by searching the physical spatial world and coming up with geometrical shapes that are inherent to the problem. These are not necessarily easy to find; but once they are formed, a much simpler computer program can be written. It is called, as may be surmised, the "geometrical approach."

The second method is more complex but is also more effective because of its flexibility. It is called the "Monte Carlo approach" and, as the name suggests, is more geared to gaming. It uses a simulation of the weapon's system as the backbone of the model, success or failure being measured by drawing random numbers.

There is one other approach that deserves mention, but with which I am not too familiar. It is called the "probabilistic approach," and a few of the important features follow.

Since every event or aspect of play is subject to uncertainties, a method can be developed to produce an uncertainty of defense. The long mathematical procedure of finding this probability turns out to be a reasonable way to conceive the real world. Figure 7 gives an example of what is meant. This is as much as I will say about this approach. The other methods will be discussed next.

Simple computer model—geometrical approach

This model's objective is to present a descriptive picture of the "goodness" of a siting. It is not intended to be a

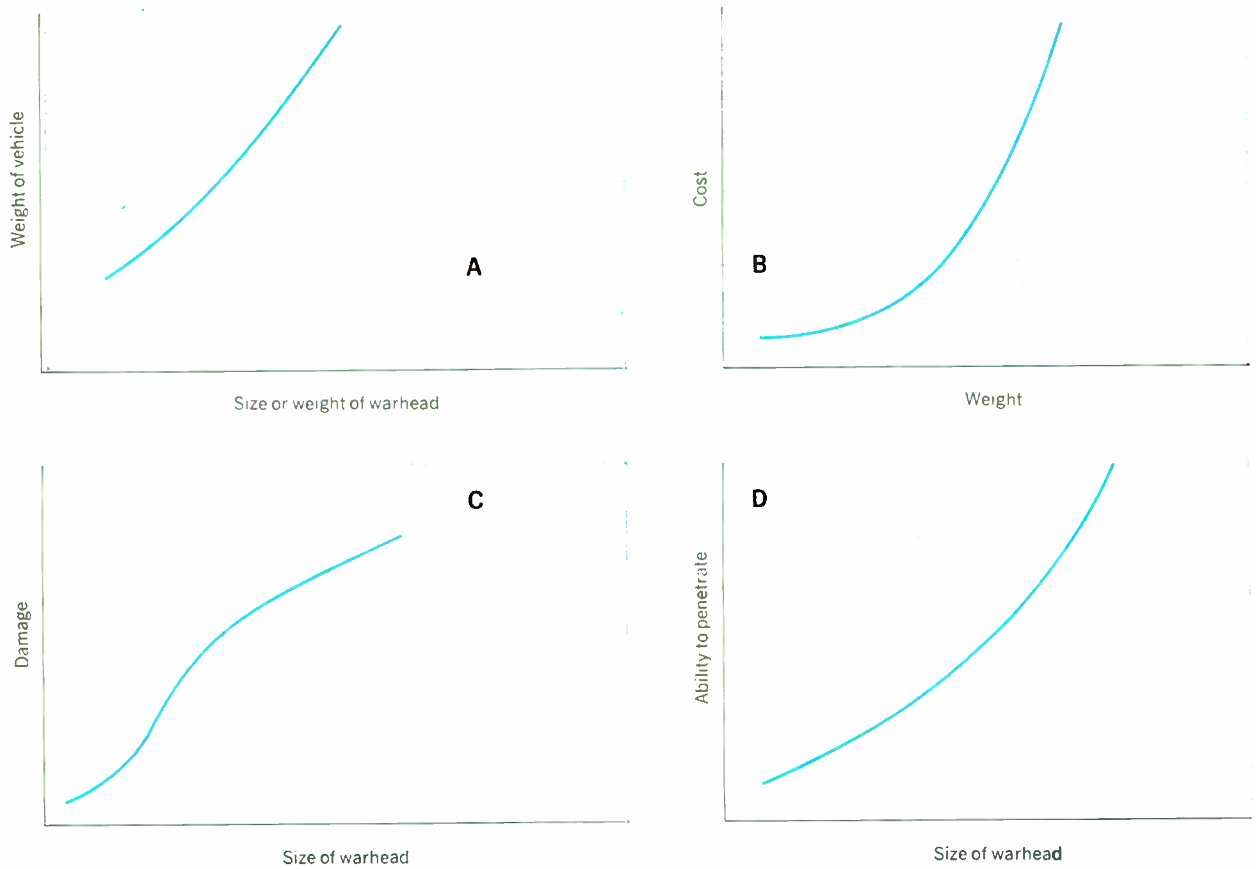
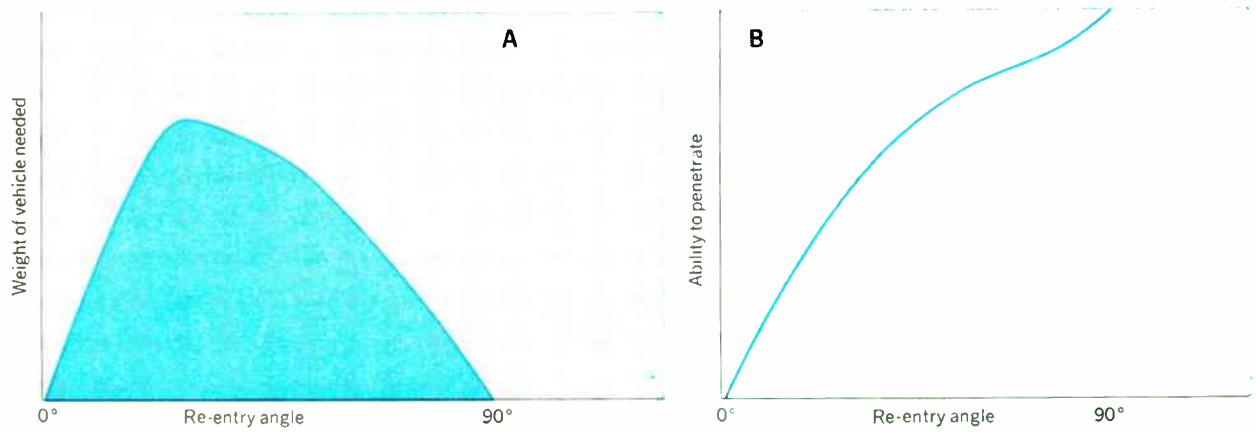


FIGURE 3. Weight, cost, and payoff differences resulting from variances in offensive warhead sizes.

FIGURE 4. Weight and penetration factors influenced by changes in missile re-entry angle.



mathematical solution. It must, however, be described by an equation; for example:

$$g_s = \sum_i \omega_i m_i \tag{3}$$

where g_s = the goodness of the site, m_i = the ith merit factor, and ω_i = weight of the ith merit factor in the overall picture. After some consideration, two merit factors can be arrived at: m—an indicator of the number of people that are not defended; and v—the volume of space that measures the imbalance of coverage of the P_w surface.

In this model, we cannot perform any simulation or any optimization analyses, since we will be looking only at the capabilities of the defense under a specific attack, given the required number of ABMs in consideration of the balanced defense approach. That is, we will assume that the unanswered questions fall out as corollaries to the theorem that if we answer the question "How well are we defending the P_w surface?" then we have reached the game-theoretic best. The goodness here is defined as a geometrical good.

Let's look at the city once more and see what restric-

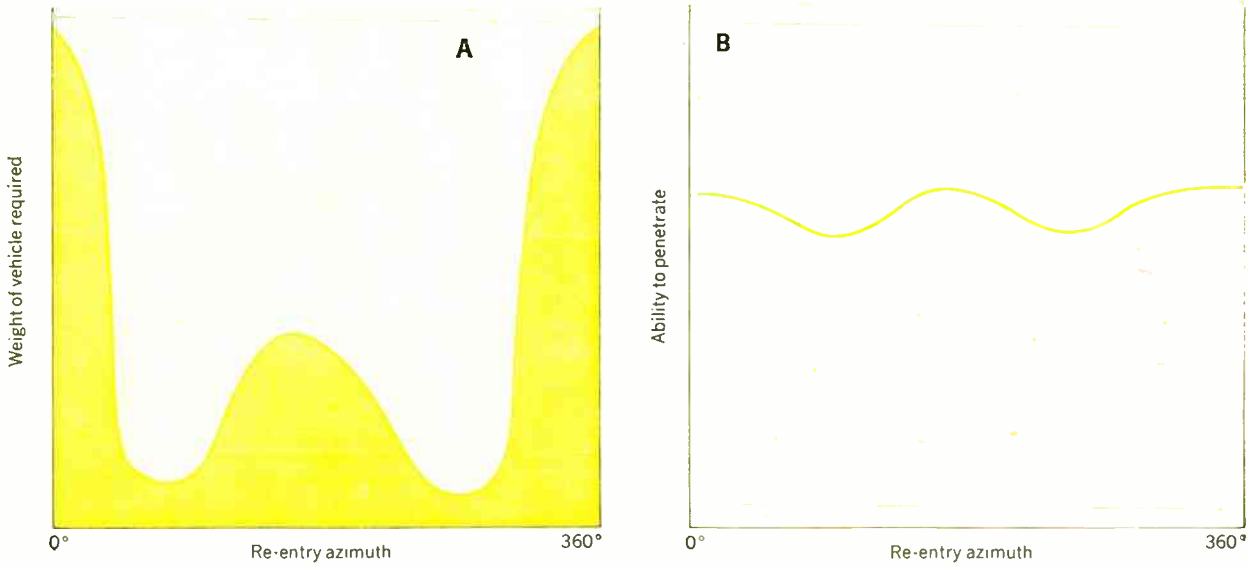
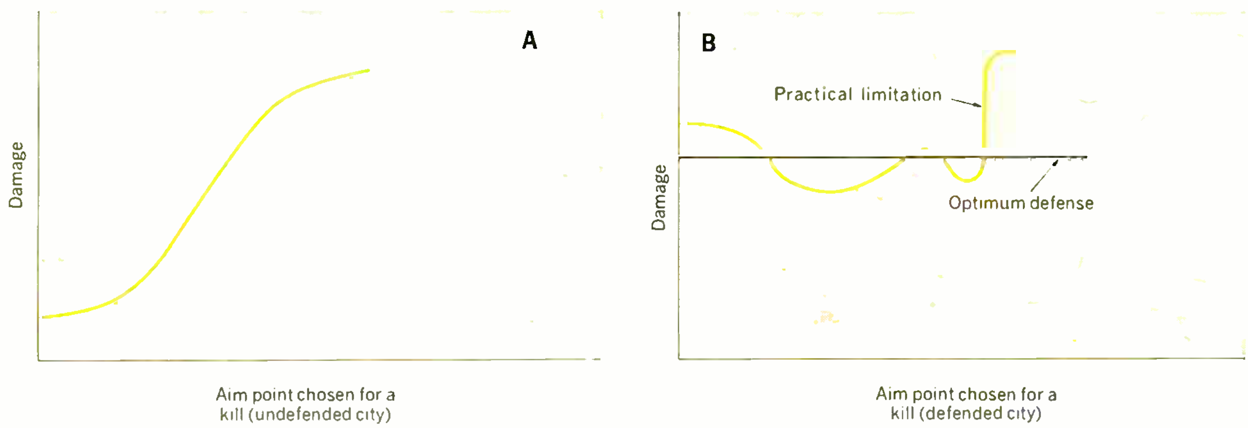


FIGURE 5. Weight and penetration factors as a function of re-entry azimuth.

FIGURE 6. Damage from missiles directed at specific aim points within a city.



tions are imposed by this intercept surface (Fig. 8).

The d_a or decision altitude, which has been set by the firing doctrine, is a measure of the discrimination and target-tracking capabilities. It says that launching the ABM before the ICBM reaches this altitude would handicap the defense by wasting missiles.

The l_a or launch altitude, which is computed as a function of the intercept point and the time required to intercept, is used to compute the surface of last penetration. Once the object has penetrated this surface, no site could launch any missiles to intercept the object and prevent its penetration of the P_w surface. A complicated expression which yields this launch altitude may be found. I mention it only for the readers who are interested.

$$l_a = h[t_o + t(i_a)] \tag{4}$$

where t_o = time required to intercept, i_a = intercept altitude, $t(i_a)$ = time function, which gives the time to impact as a function of altitude, and $h[t_o + t(i_a)]$ = altitude function (inverse of the time function), which gives altitude of object as a function of time to impact.

Note, from inspection of Fig. 8, that there is an opening or a weak spot in the defense. Note further that it is near the heart of the city. However, the opening is quickly closed if we (1) move the intercept altitude down, or (2) move the decision altitude up. The question that comes to mind then is how can we measure this weakness.

One way to measure the imbalance of coverage of P_w is by calculating the signed volume of space between the d_a and the l_a as a weighted function of d_a. Since d_a is a horizontal plane (for local defense),

$$v = \frac{\int_{d_{a_1}}^{d_{a_2}} wv \, d(d_a)}{\int_{d_{a_1}}^{d_{a_2}} w \, d(d_a)} \tag{5}$$

where v (the volume of space between l_a and d_a) = $v_o + A(d_a - d_{a_0})$, with v_o, A, and d_{a_0} constants, w = the weight associated with d_a, and d_{a_2}, d_{a_1} = the limits chosen for d_a.

Another way to measure this vulnerability is to ask how many people are affected by the hole in the defense. Unfortunately, the answer to this question is not easy to obtain by this approach. For example, by intercepting lower,

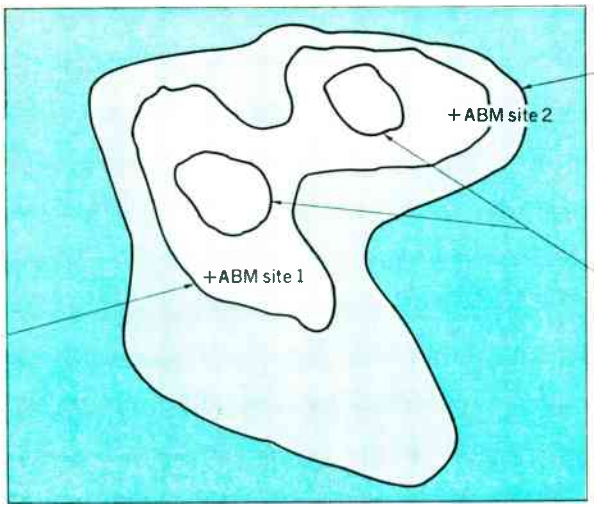


FIGURE 7. The probabilistic approach to missile defense.

one could obtain such a measure by subtracting X number of people from the balanced defense plan. However, by doing so, the entire game-theoretic balance entailed in intercepting on the P_w surface would be disturbed. About the best one can say is that, with this d_a, the siting plan fails. With others, it might work.

The formula we are looking for, then, is

$$\bar{m} = \frac{\int_{d_{a_1}}^{d_{a_2}} wm\, d(d_a)}{\int_{d_{a_1}}^{d_{a_2}} w\, d(d_a)} \tag{6}$$

where $m = 0$ or 1, depending on d_a:

$$\text{if } d_a < d_{a_0}, \text{ then } m = 0$$

$$\text{if } d_a \geqslant d_{a_0}, \text{ then } m = 1$$

with w, d_{a_1}, and d_{a_2} as previously defined.

The next step is to be able to compute these merit factors, considering all types of threats. The three variables considered will be re-entry angle α_e, re-entry azimuth θ, and weapon yield y. The formula for $\bar{v}$ then becomes:

$$\bar{v} = \frac{\iiiint wv\, dd_a\, d\alpha_e\, d\theta\, dy}{\iiiint w\, dd_a\, d\omega_e\, d\theta\, dy} \tag{7}$$

where w is considered a weight function of each variable. Equation 7 is merely the mathematical way of describing the statement: "By 'combining' all the variables together, we can look at the final merit factor and come up with our decision." This combining must be rigorously followed because (1) every statement must be verified by a certain amount of logic or mathematics, and (2) no computer program could possibly understand the layman's terms. A similar equation can be formulated for m.*

The geometrical approach is a very simple one as far as the computer is concerned. Finding the volume v for the purpose of solving Eq. (7), in principle, is straightforward. Volume is equal to area times height. Hence,

$$v = A(\bar{l}_a - d_a) \tag{8}$$

where A equals the area of the plane cut out by the defended zone, and $\bar{l}_a$ equals the average altitude of the l_a over some projection of this defended area.

In a computer program written in Fortran language, what is done is the following. Set up a two-dimensional array HEQU, describing the P_w surface. HEQU(i, j) represents the height above the surface at which P_w people would be killed if a weapon of yield y were to burst there. The point (i, j) represents a grid point, or a grid square with area A_g. If no such altitude exists for some (i, j), set HEQU(i, j) = 0.

The defended area A of Eq. (8) falls out immediately, and has the equation

$$A = \Big(\sum_{\substack{i,j \\ \mathrm{HEQU}_{i,j} \neq 0}} 1 \Big) A_g \tag{9}$$

which is easily computed. The next step is to calculate for each grid point (i, j) a launch altitude for each site. The minimum is then saved. As given before, the launch altitude is determined by knowing t_o, the time to intercept, and HEQU(i, j), the intercept altitude. The functions t and h of Eq. (4) must be fairly simple to compute or a tremendous amount of machine time can be eaten up. Hence, the best procedure is to perform all trajectory calculations (offense and defense) long before the volume calculations are done, and to save the results in the form of a set of coefficients that are, perhaps, produced by a least-squares method.

This procedure is recommended for simulations which are called event-oriented—as opposed to clock-time-oriented. The reduction in speed is usually enormous. The accuracy, however, can be somewhat limited. I will spend more time on this subject when discussing the Monte Carlo approach.

We can now see that:

*Use of the above symbolism is not intended to confuse the reader, but merely to indicate that a mathematical method exists and that, because of this, certain computer methods are necessary to solve them.

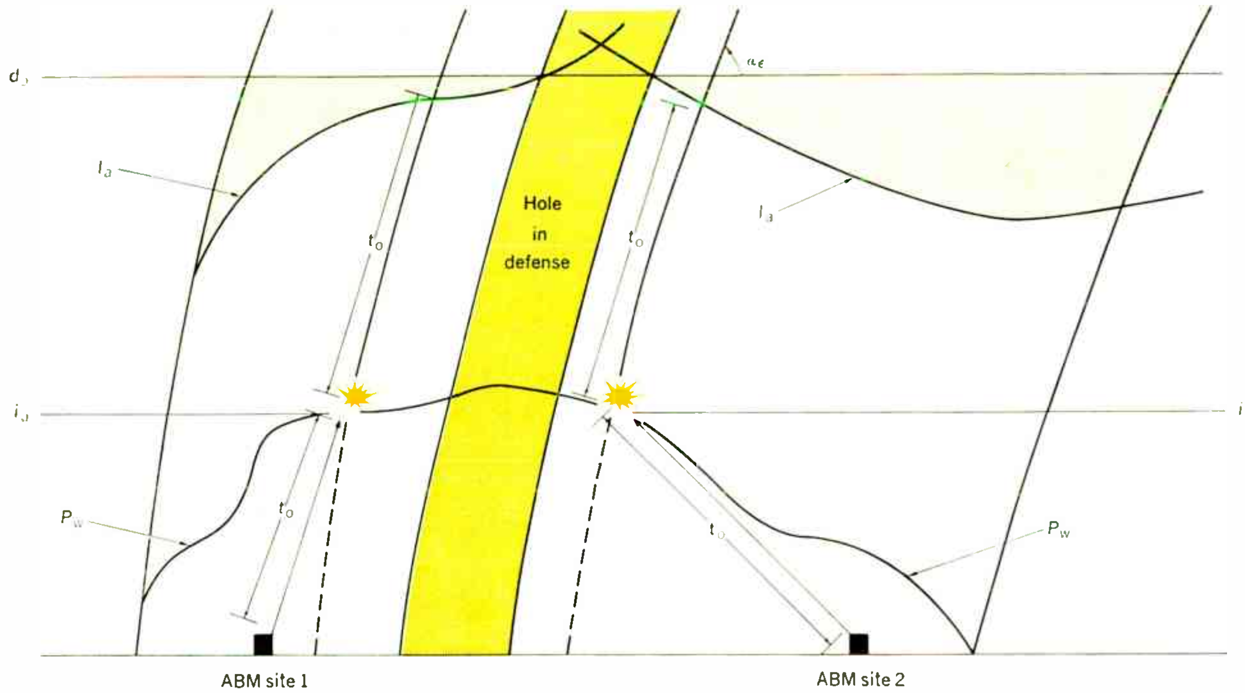


FIGURE 8. Profile of the geometrical approach to a balanced defense system. d_a = the altitude at which the defender must decide whether or not to fire, α_ϵ = the offensive missile re-entry angle, and l_a = the altitude of the attacking missile where an ABM must be launched in order to intercept on the P_w surface.

$$\bar{l}_a = \frac{1}{n} \sum_{\substack{i,j \\ \text{HEQU}_{i,j} \neq 0}} \text{minimum } l_a\text{'s} \qquad (10)$$

where

$$n = \sum_{\substack{i,j \\ \text{HEQU}_{i,j} \neq 0}} 1 \qquad \text{as in area Eq. (9)}$$

Having done this for every re-entry angle, re-entry azimuth, and weapon yield, we can obtain the answer as per Eq. (7). The number of sets of calculations that are needed of those that are time consuming (those involving missile trajectories) amounts to:

$$n_c = n_g \times n_s \times n_{\alpha_\epsilon} \times n_\theta \times n_y \times n_d \qquad (11)$$

where n_g = number of grid points such that HEQU(i, j) = 0, n_s = number of sites, n_{α_ϵ} = number of re-entry angles, n_θ = number of re-entry azimuths, n_y = number of weapon yields, and n_d = number of deployments. To achieve any kind of accuracy, the number of probable sets of calculations should be greater than n_c.

By using Eq. (11), the number of sets of calculations for a small hypothetical city would be $n_c = 1000 \times 3 \times 3 \times 3 \times 3 \times 3 = 243\ 000$ sets. To insure that we can safely solve this problem in, say, one hour of computer time, we must be able to perform each set of calculations in roughly 0.015 second. Under the foregoing scheme, this is moderately easy to do on the Control Data 3600. Programming time (assuming a background model is already developed) is minimal, perhaps 120 man-hours.

This approach can be very useful in establishing a play that will work in most cases, and that will not cost a great deal to utilize.

The Monte Carlo approach

The purpose of the Monte Carlo approach is to reduce a complicated strategy to a simple curve that can produce a merit figure. It makes use of the confidence level that is so often annoying when solving real problems. It answers satisfactorily the question: "How many people will die if I push this button?" Complicated chains of different logic can be linked together by merely playing a game in which you simulate conditions by throwing switches on or off, depending on random numbers.

The merit factor here is P_a, the price of admission, or the number of ICBMs necessary to penetrate the defense.

In the Monte Carlo approach, we are free to bring into play all manner of hazards, both offensive and defensive, which, under the geometrical approach, would have been impossible. Some of these are ABM stockpile depletion, ABM aborts, miss distance at intercept, warhead-decoy problems, warhead dispersal, and intercept geometry.

A firing doctrine or a set of rules by which one should play under every set of circumstances is required. Not all plays may be made, but the logic must be available. Here again, we maintain that we must site to defend the P_w surface, that our intercepts must be on that surface.

The Monte Carlo method must establish a simulation procedure because of the importance of the interdependence of events over a period of time, and the need to sample this interdependence at any moment in its history.

There are two types of simulation: (1) event-oriented, and (2) clock-time-oriented. Both types accomplish the same purpose but in a slightly different manner.

Event-oriented simulation is the easier of the two, since one establishes a chain of happenings in advance and determines the success or failure of the simulation by merely looking at the resultant happening.

In a time-oriented simulation we must, at each clock pulse, check for success or failure. When timing is most important (e.g., the ABM and the warhead must be at the same place at the same time and the ABM burst must be so timed), it is desirable to use this type of simulation. One will then be able to observe a truer test of a system. However, it turns out that the time-oriented simulation is good only for "feasibility studies" or "demonstration of effectiveness," and not for analysis work. It is too time consuming (on the computer), and therefore must be restricted to typical cases.

Since the case described here will be event-oriented, we shall proceed to describe the events. In this approach, we can put in as many details as we need. However, for simplicity's sake, I shall only choose five events, with meager descriptions:

1. Detection and track.
2. Assignment of site to object, computing intercept point and proper salvo.
3. Discrimination.
4. Computation of launch altitude.
5. Determination of intercept.

With regard to the fact that this approach is much more time consuming than the one described in the geometric approach, we cannot afford to have trajectories intersect every grid square of the P_w surface; therefore, we shall select various likely "aim points" in this defended zone at which the other side would aim their ICBMs. We can, perhaps, weight them, but we still must intercept on this P_w surface.

The way this method is implemented is as follows:

1. For every "threat" condition, aim points, trajectory, and warhead yield perform the next tasks.
2. Play X number of games (to be described) until you obtain a set of points corresponding to an S curve (also to be described).
3. Perform smoothing operations to obtain the S curve and compute the price of admission P_a.
4. Perform some summation to link up all the threats.

Since the purpose of a game is to obtain an S curve, I will first describe this curve. The probability of success of a "happening" often depends on the number of times the particular event takes place. This "dependency" is a function that, when mapped, traces out an S shape [see Fig. 9(B)]. The regularity of occurrence of the curve in analysis work is sufficient excuse to go into more detail.

Suppose, for example, n in Fig. 9 represents the number of rocks thrown at a snake in an attempt to kill it. The probability of actually killing the snake should increase as the number of rocks thrown increases. Further, since *any* rock could do a specified job, the curve begins to rise rapidly. However, because the snake can be only so dead, the curve has to taper off. This S curve is merely an expression of the normality of distribution of misses or failures that is well known in statistics.

The same is true if n is the number of ICBMs and p is the probability of penetration, results which our game will provide.

A game is played by selecting a particular value for n and determining, by random sampling, whether or not the offense has penetrated. By playing enough games at that level of ICBM saturation, we can establish a probability of success for that n and be able to place a point on the curve. By doing this for every n, we can obtain the desired curve. Again, unfortunately, because of time considerations, we cannot afford playing the games at every level—or, in fact, the same number of games at all levels—unless some small number of games were decided upon. Because we are only interested in a certain portion of the curve, namely, around the value that gives the price of admission, P_a, we are led to another method for calculating the points on the curve. But first a word about this important part of the curve.

From what was stated previously, we are interested in the P_a, or the number of ICBMs needed to penetrate, but we have just seen that penetration can be treated as a probability. By what means do we then determine P_a? Figure 9(B) gives the answer. We want the point on the curve where, if n goes beyond this number, the probability is very little improved and, if n goes below this number, the probability drops off quite a bit. This point is usually called the "knee" of the curve. It is obtained by sliding a line with some slope m up the p axis until it becomes tangent to the curve. The point of tangency reveals the point were P_a is obtained. The slope of the line is a function of the "cost" of extra ICBMs versus the "importance" of the penetration. The higher the slope, the costlier the missiles; the lower the slope, the higher the importance of penetration. Therefore, we want to play most of our games near the knee, the point of interest.

A well-known procedure for game playing is as follows:

1. Play a game with $n = 1$; add 1 to the number of games played at 1.

2. If you win (from a defense point of view) add $4q_n$ (integerized) to n. q_n = probability of loss at n and is assumed to be 1 at the beginning. If you lose, add 1 to the number of losses of games played at n; subtract 1 from n unless $n = 1$, in which case do not change n at all.

3. Play the game at n and add 1 to the number of games played at n; compute q_n and, unless the number of games is exhausted, go to 2.

4. When all the games are over, you typically will have a table of probabilities looking like Fig. 9(A). These data are then smoothed, resulting in Fig. 9(B). Although Figs. 9(A) and 9(B) may not resemble one another, remember that the same number of games were not played at each n.

The games involve an event-oriented simulation, and thus we will not consider such multiprocessing defense problems as radar sweeps and tracking capabilities. These multiprocessing and time-sharing problems can be implemented more in the time-oriented simulations.

This approach leads to engaging each object as it is detected. If necessary, give it a probability of detection that varies with the range of the object, and draw a random number at a critical point. Detection is important, since it enables you to determine where the intercept will be. The aim point that is selected to be bombarded is offset because of two criteria: (1) cloud pattern, and (2) "accuracy of hitting the aim point" indicator. The first has been preset by the offense, but the second must be obtained by drawing a random number that will yield an actual trajectory. The prelaunch computations are performed and a site is selected according to some rule. This leads to a computation of the launch altitude as in the geometric approach. If we have no more missiles left from any site, and if any of the remaining warheads successfully penetrates, we score a loss for this game. This last result depends on the actual aim point and reliability of the warhead.

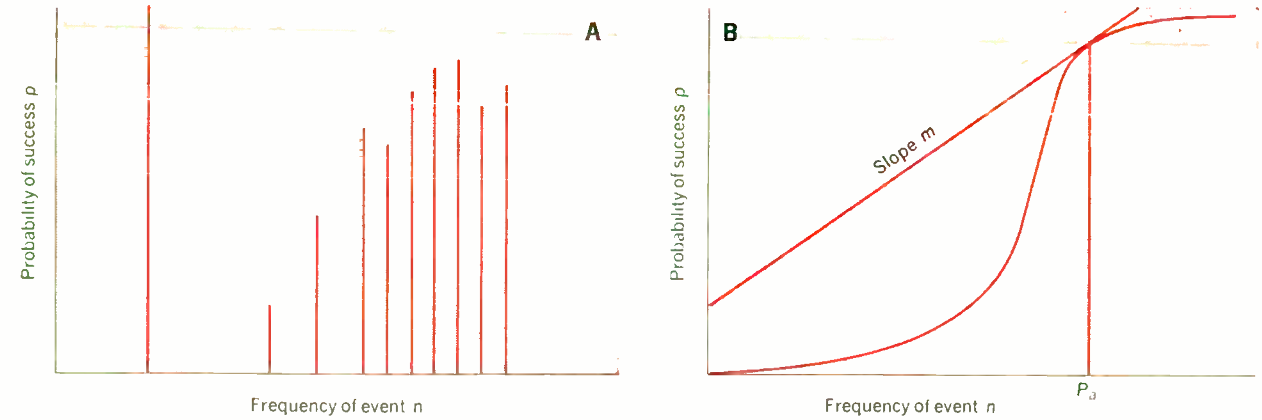


FIGURE 9. Empirical and smoothed results of a Monte Carlo approach to a balanced defense system. A—Empirical data. B—The S curve of an event-oriented simulation showing the slope method of determining the price of admission.

After establishing the assignment doctrine, we must determine whether to launch before we can discriminate, or whether we can allow the object to penetrate, assuming that it has been determined to be a decoy. We must be very certain that an assumed decoy is not a warhead before we allow it to penetrate.

Now we must insure a certain reliability of our intercept; hence we must, if necessary, fire more than one missile at the target. This number is called a salvo, and depends on the particular accuracy of intercept (or miss distance) at the intercept point (also on abort probability). This accuracy depends on the geometry of the intercept point. In our prelaunch calculations, then, we determine how many missiles we should send up (if we have them). As each missile is fired, random numbers are drawn and we determine if a success or a failure has occurred. Actually, we need not waste our time with these calculations when dealing with decoys, except in terms of stockpile depletion.

If any warhead penetrates, the game is lost. Every warhead of every ICBM is thus engaged. If all are knocked down, the game is won.

After we have determined the price of admission via playing the games as described, the summation takes place with respect to all the threats, as in the geometric approach. We then obtain our merit figure.

Computer time is much more important here since one game may take precious seconds. If, say, 50 games were played for each aspect of the threat and the same number of threats were considered as in the preceding section, the number of games actually played, n_c, is greater than $50 \times 20 \times 3 \times 3 \times 3 \times 3 \times 3$ or 243 000 games. If each game lasted one second, the number of hours of computer time would be around 67 for determining the correct deployment. This is far too much time as it turns out, and we must take away many of the niceties of the model. There can be no sophisticated salvo-size computation, nor can there be any massive site-assignment doctrine. Simple, quick, and dirty methods are best. It is not difficult to have a good simulation program that can do the siting in about 20 hours. Less time than this reduces the effectiveness of the model.

Programming time is *not* too costly. The statistical background needed to write such a program is not intense, since only simple decisions are made at each link. Such a program, once the background model has been developed, can be written in about 480 man-hours.

Summation

How can these approaches be applied to other game-theoretic applications?

The underlying motif is outlined as follows:

1. Obtain an objective that is programmable. Set up the "goodness" of the plays in such a way that an equation like the following is obtained.

$$g = \sum_i w_i m_i$$

where w_i is the weight of the ith merit m_i.

2. Determine a game-theoretic minimax that is independent of the plays that can actually be made. This is most difficult, since it usually involves some very sophisticated mathematics. This minimax will be used as the guideline for setting up the various merit factors. In other words, how well we play (according to some merit) is measured against the minimax.

3. Line up the plays on both sides so that most calculations that do not involve interaction can be performed ahead of time. When interaction does occur, perhaps a simple calculation or table search can be carried out.

4. Indicate an opponent's plays by likelihood rather than some prefabricated method; that is, establish weights for each of his plays.

5. Carry out each play to completion and obtain the final merit figures.

6. Select the play that gives the maximum g (see item 1).

Remember that computer solutions must be geared to the importance of results. These programs require much computer time at large cost; hence they cannot be overly fancy.

To write a chess-playing program, for example, that would play fairly "good" chess at a rapid rate by either of these two approaches appears to me to be reasonably difficult. The reason for this is not so much the complexity of the game, but that a chess solution is not very important. With these approaches, you pay for what you play for.

Scanning the issues

Electric Power in Finland. Inasmuch as each nation and each geographical area confronts its inhabitants with different conditions and different problems in the inevitable drift to industrialization, it is interesting to undertake comparative evaluations of the solutions to such problems. For instance, the small, rather flat, and extremely northern country of Finland, which is well known for its thousands of lakes, has drawn on those lakes as the major source of its electric power. However, these water power sources, producing more than half its electric energy, are situated rather far north of Finland's population concentrations along the southernmost and "warmest" coast.

One of Finland's special problems has been the construction of a long-distance transmission system. Thermal power, although it is rapidly gaining in Finland, has not changed the basic situation. As the economics of transmission system construction depends largely on the cost of lines, special attention was given to line design research and to system operation. This resulted at an early stage, writes Finnish engineer Erkki Voipio, in adoption of guyed towers, characterized by low weight and cost, whereas other countries have traditionally used self-supporting towers.

In Finland, too, the economy of generation and transmission is increased by a programmed regulation of water resources, the lakes playing the role of natural reservoirs. Such regulation is necessary because in the spring floods, when the accumulated ice and snow melts, there is an abundance of water power, which diminishes as the year goes on; but power consumption goes up as the water diminishes. At the end of the year, for instance, in midwinter, there are only a few hours of dim daylight in south Finland. The severe climate of Finland has also undoubtedly been a factor in the development of its severely functional architecture, probably one of that small country's most famous "exports." Even many of Finland's power stations are noted for their unusually handsome design, as is evident, for instance, in Fig. 1, which shows one power station located in the relatively unpopulated far north of Finland. (E. Voipio, "Power Generation and Transmission System of Finland," *IEEE Trans. on Power Apparatus and Systems*, January 1968.)

FIGURE 1. Pyhäkoski power station in northern Finland.

LSI Special. For those who just may not happen to know, **LSI** stands for large-scale integration, the name for the newest hot interest of the solid-state circuits community. The impact of integrated circuits on other areas of electronics is so significant, however, that the special issue of the IEEE JOURNAL OF SOLID-STATE CIRCUITS devoted to progress in LSI will appeal to a far wider audience of engineers. The progress has evidently been rapid. R. L. Petritz writes, for instance, that many of the goals that were theoretical assumptions last year are now well along the way to reality, and that the electronics industry has been relatively open with information about LSI investigations.

In his preface to the special issue, editor James D. Meindl gives us a short, broadbrush review of what has been happening in the LSI field and what the special issue is about. The interests of the solid-state circuits community, he states, have been focused on integrated circuits during the past several years. Through a potent combination of technological and economic factors, the development of digital integrated circuits has progressed at a particularly rapid rate during this period. At present, he notes, large-scale integration represents the region of singular importance in the field of digital integrated circuits. Typically, large-scale integration consists of the combination of a multiplicity of circuit elements within a basic cell and the further combination of a multiplicity of cells in a monolithic structure to form a highly complex integrated circuit or an integrated equipment component. A most significant feature of an integrated equipment component is the uniquely intimate interdependence of its material, device, circuit, and system-design considerations. The recognition of this interdependence, of the central position of LSI in electronics, and of its salient importance in solid-state circuits, provides the *raison d'être* for a special issue on the advanced work in this field.

The opening paper of the issue, "Current Status of Large-Scale Integration Technology," by R. L. Petritz, presents a rather comprehensive summary. In his introduction, he gives a useful review of terminology: LSI is seen as a system of technologies underlying the products that are called IECs (integrated electronic components).

Petritz notes that a distinction should be made between "integrated circuits" (IC) and the products, namely, "integrated electronic components" (IEC), that will result from large-scale integration technology. The principal distinction between an IC and IEC is that the latter is the result of interconnection circuits within the structure, whereas an IC is the result of interconnecting devices within the structure. Another term that is being used is MSI (medium-scale integration). Used in the context of complexity level, MSI indicates ten to 100 circuits, whereas LSI generally refers to complexity levels greater than 100 circuits.

Petritz distinguishes between "technology" and "product" for three generations of solid-state electronics. The transistor is seen as the surviving product terminology of the second generation of electronics technology. Some of its technologies are grown junction, alloy, mesa, planar, bipolar, and MOS.

Similarly, "integrated circuits" is the surviving product terminology of the third generation. Its technology terminologies include monolithic, hybrid, thin film, and thick film.

In the fourth generation of product terminology is the integrated electronic component whose technology is LSI. Petritz notes that at least three well-defined LSI technologies are under development: LSI chip technology, LSI hybrid technology, and LSI full-slice technology. And even under these technologies, there are terms such as fixed-pattern metalization, discretionary wiring, and customized wiring that Petritz indicates need precise definition. For instance, customized wiring refers to a customized metalization pattern for a specific product. In the manufacturing process, there are two distinct technologies. The fixed-pattern process uses the same metalization pattern from slice to slice in the manufacture of a specific product, whether it be a standard or a custom product. Discretionary-wiring technology is a method of enhancing yield and provides different metalization patterns for each slice in the manufacturing process, whether the product be standard or custom. A particular advantage of the discretionary-wiring technology is that for many applications it accomplishes the customized-wiring function with relative simplicity.

Petritz concludes his paper with a discussion of complexity versus cost, performance, and reliability for various LSI technologies. He considers the problem of selecting an optimum LSI technology for a given application through the use of several specific examples.

As a complement to Petritz's paper, Rex Rice in "Impact of Arrays on Digital Systems" provides a broad perspective of the effects of LSI in the design and building of digital systems. In this field, Rice says, interconnections are already the major problem. The application of LSI to digital systems, he goes on, will inevitably force the realization that interconnection design and manufacture will be more important in determining costs than all other hardware factors.

Rice concludes that to utilize LSI to maximum advantage a new set of design optimization rules must be applied. In short, a much more highly coordinated design technique that more directly couples application, language, system, and components is required to obtain the maximum benefits.

Following these broad introductory papers, C. D. Phillips and his associates deal with some of the basic design problems of integrated eqiupment components. The use of redundancy to enhance the yield of integrated equipment components is discussed by E. Tammaru and J. B. Angell. The layout of complex masks for LSI is considered by A. Weinberger, as well as by P. W. Cook and his associates, who emphasize the use of a computer for automatic artwork generation. J. E. Iwerson *et al.* discuss the application of the novel beam-lead sealed-junction technology in implementing a memory cell suitable for large-scale integration. The transmission-line behavior of the interconnections of the cells in an integrated equipment component is analyzed by I. T. Ho and S. K. Mullick. Taken together, these, and other papers in the special issue, provide a graphic description of the most recent advances in a markedly interdisciplinary field. (*IEEE Journal of Solid-State Circuits*, December 1967.)

Circuit Synthesis Breakthrough. Although the general reader may find it rather tough going, those who are involved in circuit design will be most interested in a paper in the current issue of the IEEE Transactions on Circuit Theory that represents a significant breakthrough in the synthesis of lumped-distributed circuits. The paper, "Synthesis of Finite Passive n-Ports with Prescribed Positive Real Matrices of Several Variables," is by Tosiro Koga of Kyushu University, Japan.

The idea of two variable positive real functions and matrices was first introduced in connection with the synthesis of variable-parameter networks. The importance of these functions and matrices has recently been emphasized by the success in their application to the synthesis of networks having distributed or lumped-distributed structures. However, the synthesis of these functions and matrices has been limited to the two-variable reactance case and the case where a two-variable positive real function is prescribed as a bilinear function with respect to one of the two variables. No general solution to the problem of synthesizing positive real matrices of any number of variables has been presented up till now. In his paper, Koga presents a complete solution to the problem.

Koga first shows that an arbitrarily prescribed $n \times n$ positive real matrix, symmetric or nonsymmetric, of k variables, is realizable as the impedance or admittance matrix of a finite passive n-port of those k variables. He shows further that if the matrix is symmetric, then it is realizable as a bilateral finite passive n-port.

Koga's method is said to be "noble in character" and easy to apply to any positive real rational matrix of several variables. His approach also presents a new synthesis method for the one-variable case. Furthermore, Koga shows that if the matrix is symmetric, it may always be realized without gyrators provided that extra reactances are employed. (T. Koga, "Synthesis of Finite Passive n-Ports with Prescribed Positive Real Matrices of Several Variables," *IEEE Trans. on Circuit Theory*, March 1968.)

Residual FM Noise. A state-of-the-art paper that should attract some notice examines in depth both theory and measurements of residual FM noise in two-cavity klystron oscillators used in radar systems. The study of the effects of various noise sources shows that RF beam current noise near the fundamental frequency of oscillation and near harmonics is a major source of FM. The importance of this source depends upon the value of the bunching parameter in the oscillator. Vibration and flicker effects are shown to be important sources of FM under certain conditions. The author summarizes the properties of various FM noise sources investigated. (W. R. Curtice, "Sources of Residual Frequency Modulation in X-Band Two-Cavity Klystron Oscillators," *IEEE Trans. on Electron Devices*, January 1968.)

Thin Magnetic Films. A bibliography on articles and patents in the field of thin magnetic films appears in the recent issue of IEEE Transactions on Magnetics. The listing, covering the years 1963 to 1967, supplements an earlier one published in 1964. It lists 1300 articles and 170 patents. (H. Chang and Y. S. Lin, "Bibliography of Thin Magnetic Films (1963–1967)," *IEEE Trans. on Magnetics*, December 1967.)

Positions Wanted

Classified Advertising Rates for "Positions Wanted" column: $6.00 per line per insertion. 50% discount for members of IEEE (please supply membership number with copy). Copy must be received by the 2nd of the month preceding date of issue. Address replies to Box Number given, c/o IEEE Spectrum, Advertising Department, 72 West 45th St., New York, N.Y. 10036.

Look Me Over—Electrical Engineer, 35, seeks position in northwestern USA. 8 years experience in microwaves and radar signature analysis; position need not be in these fields, however. Willing to learn your trade. MSEE from Univ. of Mich. Box 8117.

MEE, 10 yrs. cct., logic, systems design exp; some marketing, teaching. Seeks position outside U.S. Speaks German. Box 8118.

Getting full value from your computer system? Eight years experience providing and managing hardware and software technical support for fixed military command post information system users. Instructor, contract technical monitor, site manager, on-site rep. Analyzed requirements, determined needs and provided response. Available near Boston in October. MSPE BA Math 1947, BSEE Phi Kappa Phi 1958, MS Syst Engr (Reliability) Tau Beta Pi 1963. Box 8119.

B.E.E. U.S. & overseas communications/broadcast systems engineering & operations admin. exp. Seeks extended European assign. Clearances. Resume avail. Box 8120.

Faculty position in electrical engineering or computer science. PhD, 46 years old, eleven years undergraduate and graduate teaching experience, eleven years professional industrial experience. Interest in logical design, computer systems, systems engineering and energy conversion. Registered professional engineer. Prefer south. Box 8121.

EE with proven, unusual inventive ability, desires challenge of consulting work, providing new consumer or industrial product ideas to several more EE firms. Enquiries welcome. Box 8122.

EE desires faculty position in Engineering School. Interests, Computers and System Engineering. D. Sc. degree. Box 8123.

Sales Agency—power distribution & control lines, as well as lighting equipment, being evaluated by eastern Pa. & Dela. sales representatives. We proficiently manage sales thru intelligent and aggressive marketing to distributors, large industrials and utilities. Specifying agents are cultivated. For sound professional sales of your product reply to Box 8124.

PhD Engineer Manager: Educated in management, control, dynamics, electronics, heat power, machine design. Will trade in successful stable aerospace career on new beginning in large system synthesis and analysis; for example, in digitally based management, data, economic, or process control systems. Box 8125.

Technical correspondence

For the record

In the January 1968 issue of IEEE SPECTRUM, there is an article by J. Nievergelt entitled "Computers and Computing—Past, Present, and Future." This article contains an error so common in the literature that it deserves to be corrected. It is found on page 58 in the fourth paragraph.

Mr. Nievergelt claims that the designers of the electromechanical computers in the late 1930s and early 1940s "were unaware that they were rediscovering Babbage's concept." If the reader will refer to IEEE SPECTRUM for August 1964, pages 62 through 69, he will find a reprint of Howard Aiken's original 1937 memorandum that led to the development of the Harvard Mark I computer. In this original proposal, Aiken reviews some of the details of Babbage's design for both difference and analytical engines. Although it may be correct that some of the designers of the later machines were unaware of Babbage's work, Aiken was certainly aware of this previous work of Babbage.

This error is quite common in the literature and it may be worth publishing a correction.

William J. Worlton
University of California
Los Alamos, N.Mex.

Information source revealed

At the end of the article, "The Fast Fourier Transform," in the IEEE SPECTRUM issue of December 1967, a historical summary of the algorithm is given. I believe this summary was adapted from my report, "Three Fortran Programs That Perform the Cooley–Tukey Fourier Transform," Tech. Note 1967-2, Lincoln Laboratory, M.I.T., Lexington, Mass., July 28, 1967.

Norman M. Brenner
M.I.T. Lincoln Laboratory
Lexington, Mass.

Coauthor E. O. Brigham informs us that the summary, as it appeared in IEEE SPECTRUM, was adapted from notes taken during a 1967 summer course on signal processing techniques at M.I.T., and that he was not aware of the source at that time. Since then, however, he has read the report in question, and acknowledges it to be the source that his original notes were based upon.

Editor

. . . and more

Dr. Brigham has also communicated a correction of his December article, "The Fast Fourier Transform," to IEEE SPECTRUM, as follows:

$$\begin{bmatrix} S(0) \\ S(2) \\ S(1) \\ S(3) \end{bmatrix} = \begin{bmatrix} 1 & W^0 & 0 & 0 \\ 1 & W^2 & 0 & 0 \\ 0 & 0 & 1 & W^1 \\ 0 & 0 & 1 & W^3 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & W^0 & 0 \\ 0 & 1 & 0 & W^0 \\ 1 & 0 & W^2 & 0 \\ 0 & 1 & 0 & W^2 \end{bmatrix} \begin{bmatrix} x_0(0) \\ x_0(1) \\ x_0(2) \\ x_0(3) \end{bmatrix}$$

which is a corrected version of Eq. (10b), page 66.

Note that the change entails replacing the W^2 term in the second row, fourth column, of the middle matrix to W^0.

Editor

Our goof!

The IEEE SPECTRUM editorial staff regrets any inconvenience that may have been caused our readers by the inadvertent omission from the February issue of the references to D. H. Roberts' article, "Silicon Device Technology."

These references are reproduced here for those who would like to follow up on a most interesting article.

Editor

REFERENCES

1. Motorola Inc., *Integrated Circuits*. New York: McGraw-Hill, 1965.
2. *Philips Tech. Rev.*, vol. 27, no. 7, 1966.
3. Hester, D. R., "Processing monolithic silicon integrated circuits," *Electron. Power*, pp. 323–327, Sept. 1966.
4. Hester, D. R., "Organization of silicon integrated circuit development," *Microelectron. Reliability*, vol. 7, pp. 95–104, May 1967.
5. Foss, R. C., *Proc. 2nd Oxford Electron. Lecture Course*, 1967.
6. Fairfield, J. M., and Schwuttke, G. H., "Precipitation effects in diffused transistor structures," *J. Appl. Phys.*, vol. 37, pp. 1536–1541, Mar. 15, 1966.
7. Williams, B. D., Private communication.
8. Williams, B. D., Patent applied for.

9. *Electronics*, vol. 38, p. 125, Jan. 11, 1965.

10. Allison, D. F., "Dielectric isolation for integrated circuits," *Solid State Design*, pp. 29–34, Jan. 1965.

11. Brothers, J. S., and Gay, M. J., "Monolithic integrated circuit capacitors," *Proc. IEEE (Correspondence)*, vol. 53, p. 644, June 1965.

12. "Integrated electronics: Design ingenuity is the key to success," *Electronics*, pp. 100–106, Aug. 7, 1967.

13. Manasevit, H. M., and Simpson, W. I., "Single-crystal silicon on a sapphire substrate," *J. Appl. Phys. (Communications)*, vol. 35, Apr. 1964, pp. 1349–1351.

14. Allison, J. F., Burns, J. R., and Heiman, F. P., paper presented at the International Solid-State Circuits Conference, Philadelphia, Pa., Feb. 15–17, 1967.

15. Green, B. J., "The design of active filters for hybrid integrated realisation," *Proc. Conf. on Integrated Circuits*, Eastbourne, England, May 2–4, 1967, IEE (London) Conf. Pub. No. 30, pp. 76–90, 1967.

16. Preston, K., Jr., "Computing at the speed of light," *Electronics*, vol. 38, pp. 72–83, Sept. 6, 1965.

17. Noble, P. J. W., "Light sensitive arrays based on photodiodes combined with MOS devices," *Proc. Conf. on Integrated Circuits*, Eastbourne, England, May 2–4, 1967, IEE (London) Conf. Pub. No. 30, pp. 251–257, 1967.

18. Atalla, M. M., paper presented at the Solid State Devices Conference, Manchester, England, Sept. 5–8, 1967.

19. Becke, H. W., and White, J. P., "Gallium arsenide insulated gate field effect transistors," *Proc. Internat'l Symp. on Gallium Arsenide*, Reading, England, Sept. 1966. London: Institute of Physics and Physical Society, 1967, pp. 219–227.

20. Turner, J. M., "Gallium arsenide field effect transistors," *Proc. Internat'l Symp. on Gallium Arsenide*, Reading, England, Sept. 1966. London: Institute of Physics and Physical Society, 1967, pp. 213–218.

Better applications needed?

The article, "Negative Conductance in Semiconductors" by Herbert Kroemer, which appears in the January 1968 issue of IEEE SPECTRUM, is a long-awaited paper, that reveals in a very interesting way the important aspects of negative conductance in semiconductors and their uses at microwave frequencies and beyond.

Although it answers several valuable questions, it also creates new ones. The article is limited to two main areas; namely, the specific phenomena of (1) the avalanche transit-time effect and (2) the Gunn effect.

Since the middle fifties we have been using negative-conductance devices such as Shockley diodes and double-base diodes or unijunction transistors, but in rather limited areas.

In the early 1960s, the Esaki diode, with all of its attractive features, was also limited in its use, even more than the two previous ones. It seems that the problem lies, not in the invention and development of new negative-conductance devices, but rather in the effective use of such devices by the everyday practicing engineer, especially at microwave and higher frequencies.

Limitations to their use stem from several factors. For example, the Shockley diode is a slow device because of its structure and not because of its avalanching action. No matter what kind of geometry one uses, we have to face the speed limitations of three junctions and the transit time within the device.

Second, fabrication is still not under full control.

Third, the device is bilateral between control element and output, as well as input. The unijunction transistor has similar limitations and a high carrier delay between the base contacts.

The Esaki diode, which is suitable for small-signal amplification, is also highly bilateral. Moreover, it is very difficult to stabilize it in its negative-conductance region. One must use circulators to make it unilateral, and even this technique is good only for a single frequency.

No doubt, it is relieving to see that more negative-conductance devices are at the disposal of scientists and engineers, but some conservatism regarding how these devices can be effectively utilized in small-signal amplification or switching is imposed. To insert a third terminal or second junction to a two-terminal device will reduce its frequency response. Therefore, the new device or phenomenon will have as limited an application as its counterparts, and the only promising area in microwaves would be its use as an oscillator.

It would be appropriate to mention at this point the two important concepts and factors that will accelerate the effective use of negative-conductance devices. One is the study of such devices as functional elements,[1,2] giving the electrical engineer a chance to study the inherent physical behavior of semiconductors; second, the creation of nonmetallurgical space-charge-layer regions within the structure may reduce the transit-time distance between emitting and collection areas due to the elimination of a mechanical contact area. For example, the electric field associated with the gate of the MOSFET device establishes a space-charge layer or channel region between the two contacts. Although present-day MOSFET devices have high time constants, their structure gives a clue as to how the high-frequency devices can be fabricated and utilized without having mechanical contacts.

The electric field, which is the result of ionization, can be realized without making a mechanical contact on the control grid of the device. Thus, the distance between source and drain can be reduced to subnanosecond transition times, and low-frequency noise, due to interface states, will be eliminated.

It can be said that, although the article is of great interest from a physical, as well as a circuit, point of view, it leaves the reader with several important unanswered questions. Thus, it would be appropriate at this point to see an article that covers the phenomena of negative conductance for all existing devices, so that we can know where we stand in our skill at creating negative conductance, and in making use of such devices effectively.

For example, the increase of collector current with reduced voltage across the collector of a transistor is still an argument if it stems from a negative conductance created within the space-charge-layer region of the transistor.

Vasil Uzunoglu
Cockeysville, Md.

REFERENCES

1. Uzunoglu, V., "Some future aspects of microelectronics," *Wescon Conv. Record*, Session 5, 1966.
2. Uzunoglu, V., "Solid-state opportunities," *IEEE Spectrum* (*Tech. Corr.*), vol. 4, p. 160, July 1967.

Science a tool?

Your November 1967 IEEE SPECTRUM "Spectral Lines" requires a reply. Some basic researchers possibly have little idea how their results may be used. However, these are balanced by other narrow specialists working at the applications; and the overall program is usually guided by applications- and mission-oriented managers.

Every person and organization is inhibited by "financial, organizational, and environmental considerations." The basic cost-effectiveness concepts are generalizations of inescapable mass-energy conservation laws, which are not dismissed by the wave of a pen. All public funds should be regulated carefully to show optimum cost effectiveness in accomplishing public goals. Also, good cost-effectiveness studies can be applied to projects where benefits are "a generation away." Science should never "be supported as *science*" alone, because it is a tool—a means, not an end.

Also, in what units and by what studies has "the effectiveness of pure research . . . been amply demonstrated"?

Robert L. Hamson
Malvern, Pa.

Book reviews

Laser Receivers—Devices, Techniques, Systems, Monte Ross—*John Wiley & Sons, Inc., 605 Third Ave., New York, N.Y., 1966; 385 pages, illus., $14.95.* This book is a comprehensive survey of laser communication systems and so, in spite of its relatively modest length, goes well beyond the scope of its title. The coverage is, of course, not exhaustive of every aspect of this very wide subject, but the author has managed to deal with most of the important topics involved. As in other laser applications, optical communication problems involve some techniques that are of common knowledge among engineers, but also others that are neither familiar nor yet very well established on their own ground. It is in such situations that books of the present type can make a significant contribution by presenting material from several areas in a way that emphasizes coordination rather than complete coverage.

The nine chapters (excluding the introduction and a concluding chapter on the future of information systems based on lasers) can be divided into two groups. Four chapters devoted to receiving techniques and devices, modulation techniques, and system considerations and configuration, cover the principal subject of the book. The other five chapters discuss background material and supporting techniques: noise considerations, information theory, background noise, transmitting media, and optical components.

The presentation is good and draws on material from many recent papers, listed as references at the end of each chapter. Its value for actual design purposes is enhanced by the inclusion of numerous curves and tables. The work is aimed at engineers with an electronics background and little acquaintance with optics and quantum electronics; such readers will find this book most useful for an integrated survey of the field as well as for a practical introduction to actual evaluation and design of devices. In an area as fast moving as laser techniques, a great deal of the material will certainly have to be rewritten if future editions are published. However, the essential objectives of this work are not dependent on current devices, as they consist of the coordination of material now scattered in many publications pertaining to different specialties.

J. R. Fontana
University of California
Santa Barbara, Calif.

Quantum Optics and Electronics, Les Houches 1964 (in French and English), C. DeWitt, A. Blandin, and C. Cohen-Tannoudji, eds.—*Gordon and Breach Science Publishers, 150 Fifth Ave., New York, N.Y., 1965; 620 pages, illus., $10.50, $8.50 ppbk.* This book contains the lectures delivered at Les Houches during the 1964 session of the Summer School of Theoretical Physics of the University of Grenoble. It comprises eight series of lectures that range from a basic formulation of the quantum theory of radiation to specialized topics such as line-width theory and nonlinear optics, delivered by leading men in the field from France and the United States. This does not result in an easy introduction to the subject, but it is a thorough and reliable coverage of the principles involved in stimulated-emission processes.

In his survey of quantum theory applied to electromagnetic fields and interactions, N. M. Kroll treats the various processes relating radiation to classical sources and quantized systems, and explores interactions involving nonlinearities, scattering processes, and energy-level shifts. Dealing with optical coherence and photon statistics, R. J. Glauber discusses one-atom and n-atom photon detectors, correlation functions, the interpretation of intensity interferometer experiments, phase space distributions, and related topics. One of the most interesting portions of the book is J. Brossel's account of optical pumping theory and experiments. The work carried out by Brossel and his colleagues and inspired by A. Kastler at the Ecole Normale of Paris. The Nobel Prize awarded to Kastler is a measure of the importance of the experiments for the understanding of radiative interactions.

In his lectures, W. E. Lamb presents his theory of optical masers along the same lines used in his June 1964 paper in the *Physical Review*. There is also a preliminary discussion and additional material on line width. N. Bloembergen, lecturing on nonlinear optics, gives a detailed mathematical treatment, but manages to enliven some of the rather forbidding portions of the theory by physical interpretations and simplified models. It is a good summary of a considerable amount of work. There are also lectures on particular types of lasers by A. Javan and P. Aigrain, who describe gas and solid-state devices, respectively, and a series on line width by J. M. Winter.

In the three years since the lectures were delivered, much progress has been made in quantum electronic devices. However, the material of this book has not become obsolete or less significant, because it deals primarily with the fundamental processes involved and not with detailed descriptions of particular applications. In fact, it is still about the best and most comprehensive collection of articles on basic quantum optics to be found between book covers.

J. R. Fontana
University of California
Santa Barbara, Calif.

Progress in Biocybernetics, Vol. 2, Norbert Wiener and J. P. Schade, eds.—*American Elsevier Publishing Co., Inc. 52 Vanderbilt Ave., New York, N.Y., 1965; 263 pages, illus., $14.50.* Early in 1964 Norbert Weiner went to Amsterdam as visiting professor in cybernetics at the Central Institute for Brain Research. He participated in the plans for this volume, but died before it could be assembled. Accordingly, what was intended to be a collection celebrating his 70th birthday became a tribute to his memory. One cannot help feeling that he deserved better.

Progress in Biocybernetics, Vol. 2, like too many other compilations of invited papers, is very uneven. Dennis Gabor contributes a very short (nine-page) discussion of two-dimensional image processing. Even in these few pages one feels again the power of Dr. Gabor's imagination and geometric intuition. Unfortunately such a broad-brush treatment can hardly do the subject justice; on reading this essay the casual reader would never find out that filtering and information compression in visual displays is a very active field of study. By contrast, Newell and Simon contribute a long (57-page) discussion of a human performance in chess playing with most of the discussion referring to a single game. The authors have long held to the goal of understanding human