SESSION X SUPERCOMPUTING IN PSYCHOLOGY

Bert F. Green, *President*

Advanced computing in psychology

BERT F. GREEN

Johns Hopkins University, Baltimore, Maryland

CYNTHIA H. NULL

Federation of Behavioral, Psychological and Cognitive Sciences, Washington, DC

GEORGE FURNAS

Bell Communications Research, Morristown, New Jersey

MARGARET HAGEN

Boston University, Boston, Massachusetts

and

DAVID RUMELHART

University of California at San Diego, La Jolla, California

In 1984, the National Science Foundation began a large initiative to provide the scientific community with access to Class VI computers. With the support of the NSF Program on Memory and Cognitive Processes, the Federation of Behavioral, Psychological and Cognitive Sciences invited a group of scientists to a workshop that considered the advanced computational needs of several areas of psychological, behavioral, and cognitive sciences. This paper presents a summary of the conference deliberations.

Psychology has a long history of computer use. By 1960, many investigators were using computers in such research endeavors as on-line experimental control, simulation of mathematical models, artificial intelligence (AI), and data analysis (Green, 1957; Newell, Shaw, & Simon, 1958; Wrigley, 1957). Today, computer use is ubiquitous. Computers are used for stimulus generation, experimental control, data collection, data analysis, and designing and testing information processing models of behavior. Psychologists are also interested in the design of user-friendly computer systems and special computer tools to aid human-computer interaction.

Many of these applications now encounter computation problems, problems whose scale must be drastically reduced in order to be feasible on the available computing machinery. The National Science Foundation (NSF) initiative for scientific computing thus comes at a time critical for advancement in many areas of behavioral science.

The core of the NSF initiative, now embodied in the NSF Office of Advanced Scientific Computing, is the pro-

vision of access to one of several computing facilities offering Class VI computer service (see Appendix B). Other expressed interests of the initiative include the exploration of alternative computer architectures and the development of computer networks, to facilitate access to special resources such as Class VI computers and to foster communication among scientists working at different institutions.

With the support of the NSF Program on Memory and Cognitive Processes, the Federation of Behavioral, Psychological and Cognitive Sciences invited a group of scientists to a workshop that considered the computational needs of several areas of psychological, behavioral, and cognitive sciences (hereinafter collectively called psychology). Particular attention was placed on the ways in which psychologists could make important use of Class VI computers.

Class VI computers, such as the Cray 1S and the Cyber 205, are larger and faster than current mainframe computers like the IBM 3033. The combination of memory size and processor speed provides from 20 to 50 times the computing power of current mainframes. In addition, Class VI computers include "pipeline" arithmetic units, capable of extremely efficient computation of vector arithmetic. Programs carefully designed to take full advantage of the vector arithmetic units can gain another order of

Reprint requests should be sent to Cynthia H. Null, Executive Director, Federation of Behavioral, Psychological and Cognitive Sciences, 1200 Seventeenth Street N.W., Washington DC 20036. Partial support for this project was provided by the National Science Foundation, Grant BNS-8414453. Special thanks go to all those who attended the workshop.

magnitude improvement in computing power. The Class VI machines are FORTRAN-oriented, but their general speed and size make them useful for all computation-intensive projects. (A brief introduction to Class VI computers appears in Appendix A.)

This report identifies areas of psychology in which advanced computer resources are especially needed: cognitive science, including artificial intelligence and information processing models, psychometrics and statistical data analysis, and computer graphics and visual perception. Some applications, especially in data analysis, can make excellent use of current Class VI supercomputers. Other applications need a more flexible computer, offering large-scale parallel processing. The report also considers some problems in the use of remote supercomputers, including human-computer interface and computer networks. A need for networks of cognitive scientists working cooperatively is also identified.

Despite this broad spectrum of uses, the report does not treat computer use in some areas of psychology, including psychobiology, neuroscience, sensory psychology, and physiology.

COGNITIVE SCIENCE

Cognitive science is an interdisciplinary effort to construct models of mental phenomena. It draws upon findings in psychology, linguistics, computer science, philosophy, and neuroscience. Recent progress in this very active field has been reviewed in a report to the President's science advisor (Newell & Estes, 1983), which contains further details.

Cognitive science research depends heavily upon the use of computers to simulate models of mental processes. Two examples will illustrate the utility of this work. Computational models of vision greatly aid in our understanding of the links between visual sensation and the perception of complex objects. At another level, computer models of verbal processing address the question of how educated people read words so easily. How do people rapidly identify the meaning of a visual symbol? Perhaps more interesting, why is it that people who have large vocabularies are quicker at recognizing simple words? What sort of information retrieval system gets better as its data base increases? Substantial advances in our understanding of word recognition are made by exploring computational parallels of the process. Such work is impossible without the computer.

This section of the report addresses two questions. First, a class of cognitive science problems is described that could be attacked using Class VI computers as presently designed. Then, several cognitive science problems are described that require major computing support on machines with other architectures.

Current Application With Immediate Need for Vector-Based Supercomputers

Human perception and cognition are complex behaviors. Does the mechanism that produces them have

to be similarly complex? Recently, a number of models have been proposed to account for aspects of perception and cognition by the interaction of a large number of simple (neuron-like) computational elements. These models, inspired by the neural architecture of the brain, utilize algorithms that are computationally intensive, can be characterized as semilinear, and are similar to those used to model fluid dynamics, spin glasses, and other physical systems. Specific topics that have been addressed by this approach include depth perception, perception of printed and spoken language, and learning and memory. This approach to perception and cognition could be enhanced immediately and substantially advanced by ready access to high-speed vector-based machines.

Limitations of Current Technology

Research in cognitive science has been limited by the nonavailability of adequate computing facilities. The models that can be investigated are too small in scale to be realistic models of either perception or cognition. As a result, existing models focus on a single level or aspect of a much larger cognitive system. The effective behavior of each component of cognition can be only partially understood by considering it in isolation.

In addition, recent conceptual advances (Ackley, Hinton & Sejnowski, 1985) have demonstrated that introducing randomness into models of cognition can actually improve their performance. However, the introduction of randomness requires repeated Monte Carlo simulations to evaluate the behavior of a particular model.

Models that account for the gradual acquisition of perceptual and cognitive skills generally postulate the modulation of the strengths of the connection among the processing elements. Simulation requires taking the model through the processing of an entire ensemble of stimuli many times.

Most models have parameters that must be estimated by assessing the models' behavior for many different sets of parameter values. Finding the optimal set of parameters can require staggering numbers of simulations.

Current simple models now run slowly on VAXes (about 1,000 times as slowly as the real-time processes they attempt to simulate). Any one of the scientifically critical extensions described above would slow down the acquisition of results by at least another two orders of magnitude. Obviously, any increase in speed of two or three orders of magnitude, as may be possible in the existing Class VI computers, would promote substantial progress in these areas. Such progress could assist in the development of a computationally and psychologically adequate model of reading, learning, and visual and auditory perception.

Symbolic Processing Applications

Many modeling efforts in cognitive science involve the manipulation of symbols and are not arithmetic. Symbol manipulation would benefit from a special architecture not found in present Class VI machines. This section will present some examples of symbolic computation and de-

scribe briefly the advanced scientific computational facilities required for this research.

Language is the prototype example of a symbolic system. Detailed models now exist that simulate processes of perception, attention, decision, and storage and retrieval of information in memory (e.g., Anderson, 1983). These models have been evaluated and modified by comparing their performance with substantial bodies of detailed experimental data obtained from human performance.

In addition, many advances have been made in the study of problem solving. The analysis of human problem solving using concepts of information processing began in the early 1970s. Detailed models have been developed that simulate problem-solving performance in several technical domains, including detailed simulation of patterns of errors, knowledge required for understanding problem situations, and knowledge required for strategic planning of problem-solving activities (e.g., Newell & Simon, 1972). Most realistic models in this class involve very large discrete knowledge bases that are searched and interrogated. Integrated models that include processes of representation, reasoning, and problem solving must be constructed in order to address important theoretical, as well as practical issues. But a factor limiting the rate of progress in these research areas is the shortage of large-scale computational resources, especially resources for symbolic processing. The amount of computation required, the amount of storage, and the need for truly parallel computation are substantial.

The Cognitive Science Tradition of Computation

Understanding the computational needs of cognitive science in relation to those of the physical sciences involves bridging two computational traditions that have been separate for at least 25 years. The two traditions developed from concerns for numerical versus symbolic processing. The divergence began in the late 1950s and early 1960s with the development of the early symbol processing languages such as IPL and LISP. Whereas one community focused on the development of higher level languages capable of expressing the kinds of symbol processing characteristic of intelligent systems, the other community continued to use FORTRAN and sought ever faster number crunchers. On the one hand, the symbol processing tradition was instrumental in the development of interactive computational systems. Interactive processing led to the development of rich programming environments such as INTERLISP, which, although computationally expensive, decreased the production cycle and thereby increased productivity. This was followed by the development of the workstation concept, including the use of "windows," spatial input devices like the "mouse," and bit-mapped graphic displays. The SYMBOLICS 3600, a system microcoded for LISP processing with many special hardware and software facilities, is an example of development in this tradition. Development in the numeric

tradition, on the other hand, has focused on optimizing software and hardware for the fastest, most efficient execution of floating-point arithmetic problems. The vector-based Cray XMP exemplifies such developments.

Summary

Some cognitive modeling is computationally intensive. The problems psychologists are attempting to solve are at least as complex as those of the natural sciences. Cognitive scientists need computer facilities that are several orders of magnitude faster and several times larger in memory storage capacity than those available today. Some problems in this area are essentially numerical, involving very large floating-point computations, and for these the vector-based machines are well designed. Appropriate access to these machines can offer substantial improvement immediately, by enabling work on problems of a scale not otherwise possible.

Another class of computations, best characterized as symbolic computations, is not well suited to the architecture of the vector machines. Some current machines are designed to optimize symbolic processing, but they are not sufficiently fast or efficient. Supercomputers of a different architectural design seem to be required for these problems.

Some problems are amenable to parallel computational algorithms, and, therefore, it seems important that parallel computational facilities be made available as well. In addition, since some computational tasks could best be solved through the development of special-purpose hardware, facilities for the rapid development of VLSI hardware for specific applications should be made available.

PSYCHOMETRICS AND STATISTICAL DATA ANALYSIS

Quantitative and statistical methods have had profound impact on many, though not all, fields of psychology since the early foundations of psychological science over a century ago. Indeed, some of the earliest traditions in psychology were concerned with such issues as the quantitative specification of the relationship between physical stimulation and psychic responses (psychophysics) and statistical relationships between mental test scores and performance in practical situations such as school and military service (psychometrics). These historical traditions have made the development and testing of psychological theories intimately related to the growth of mathematical and statistical methods. Today, psychology is one of the great consumers as well as developers of applied mathematical and statistical techniques for effectively testing theories, evaluating data, and giving insight into important psychological, behavioral, and social processes.

The nature of psychometrics and statistics integral to the growth of psychological science can be seen in experimental as well as nonexperimental domains. Virtually all subdisciplines within psychology that rely heavily upon the experimental method to evaluate theoretical proposi-

tions (such as experimental psychology, market research, psycholinguistics, etc.) rely extensively upon statistical methods to evaluate their hypotheses against the null model of chance. A number of fields such as clinical, developmental, and social psychology carry out substantial parts of their research on naturally occurring populations and must use nonexperimental or quasiexperimental methods to evaluate theoretical propositions.

The general linear methods involving structures with latent variables have been developed to deal with such data. These methods represent the convergence of research traditions in psychometrics (factor analysis), econometrics (simultaneous equations), and biometrics (path analysis) and are particularly important because they are able to disentangle theoretically meaningful influences of constructs on each other from the relatively uninteresting effects of random errors of measurement. As a consequence, theories can be tested with nonexperimental data using modern methods of multivariate analysis.

However, these methods are heavily computation-bound. For example, the extensive use of the program LISREL in recent years has "broken the bank" of many departments. The nonlinear optimization involved is based upon problem sizes that can far exceed the typical requirements in engineering or the physical sciences. Nonetheless, many researchers appear unable to obtain the needed resources to carry out the work that is at the forefront of psychological science.

Methods of data analysis and models of individual performance that press the computer's capacity generally involve iterative vector computations for fitting model parameters to large data sets. Many involve large eigenvalue problems, themselves iterative, as steps within a single iteration of the large problem. Thus, these problems are well suited for the Class VI computer. The major types of analysis currently experiencing computer overload are covariance structure models, log-linear models of categorical data, item response theory of measurement, non-metric multidimensional scaling, and special models in clustering and classification.

COMPUTER GRAPHICS AND VISUAL PERCEPTION

Controlled laboratory studies of the perceptual process depend upon computer-generated images. The perceptual process involves the apprehension of environmental information or energy that specifies or covaries with the size, shape, distance, location, orientation, color, and composition of objects and events. The information manifold is complex, and successful understanding of the specifics of perception depends on the ability to isolate, control, and systematically manipulate sources of information and types of energy.

It is extraordinarily difficult to alter systematically the real-world environment. It is possible to create and manipulate environments in miniature, but the effort is self-limiting. For these reasons, the stimulus displays com-

mon in traditional visual research are extremely impoverished. Film has helped ameliorate these problems, but, again, the variety and complexity of possible manipulations are constrained by the photographic medium. Film creates only a visual representation of the visual information, not a formal description of it. In applications in which it is essential to combine formal descriptions of information with visual displays, computer specification and generation of images are the only feasible alternative. In applications in which systematic environmental manipulation is simply not logistically feasible, simulation is the only cost-effective approach.

Visual perceptual research has a long history of advances correlated with progress in computer graphics. With today's improved state of the art in graphics, vision scientists have the opportunity, never available before, to advance their science. In various local installations, researchers seeking to exploit the potential of computer graphics for visual research are often investing considerable time in the development or modification of special-purpose algorithms.

In several problem areas, some modest success has been realized, at least where the required images have been fairly simple or few in number or where realism was not necessary.

As local facilities improve, so too will the variety of domains and applications employing computer-generated images in perceptual research. Certain applications can be successfully implemented with currently available resources, as long as stimulus complexity is minimal, the number of images is small, and the presence of artifacts and mediocre photo realism is tolerable. Not every research question needs to be addressed with thousands of high-resolution naturalistic images. But certain parameters intrinsic to the field constrain and direct that growth. For example, three characteristics of a computer-generated image—resolution, scene complexity, and degree of freedom from artifacts—determine the feasibility of creating adequate stimulus displays with the hardware and software available to the moderately well-funded researcher.

It has become increasingly apparent, however, that, for certain research areas in the forefront of modern psychophysics, such constraints on the quality and number of the stimulus displays are simply intolerable. In these areas, images must be realistic and/or produced in great number, but the time required for stimulus generation rises prohibitively. A high-resolution film (4,000 × 4,000 pixels) requires 10 h of computation per frame on our IBM 3081D, or 8.2 years for a 10-min film. A Cray could do the job in a few days.

No research scientist can wait 8 years to create a stimulus display for one experiment. Research applications requiring stimuli of this complexity are simply not possible without the aid of a supercomputer. Given today's technology, the example above may seem an extreme case representing an unnecessary ideal in stimulus quality and complexity. But there are numerous and increasing appli-

cations of more modest demands that cannot be realized without significant and often crippling consequences for quality of research. Conspicuous among such applications are the investigations of motion perception and of spectral and achromatic surface reflectance. In some paradigms, the quality of the information itself is being investigated; in others, a fine-tuned analysis of subject response is the issue. In both cases, the investigations are seriously hampered by lack of access to a supercomputer.

Summary

Supercomputers will significantly advance the frontiers of visual perception. Motion perception experiments that are absolutely dependent on generating a large number of images can be made feasible finally through the power of the supercomputer. Also, research on surface microstructure such as shading, texture, and color, for which a high degree of parametric specification is crucial, is made considerably more feasible with the power of the supercomputer. Finally, for the applications described above, for which high resolution, freedom from artifacts, a significant number of images, and/or the violation of real-world constraints is needed, research will be significantly advanced by access to image generation on supercomputers.

It should be noted that, depending on the application desired, maximal realization of supercomputer access is variously dependent on the development of graphics facilities in local psychophysical laboratories, the development of common software resources, access to network(s), and collaboration between visual science and computer graphics specialists.

RESEARCH IN HUMAN-COMPUTER INTERACTION

The study of cognitive aspects of computer use is a new field in psychology. As such, only a few areas in the study of human-computer interaction have progressed to the point where they push the limits of our available computational resources. However, with the rapid increase of human-computer interaction research, several more areas can be expected to suffer from computational limitations.

Researchers in human-computer interfaces are confronted with a very rich and largely unknown problem space. As a result, workers are often forced into an exploratory and inductive mode to identify cohesive research problems and identify relevant variables. This exploratory mode of research depends on a methodology of rapid prototyping, in which various interface designs are developed and evaluated using a repeating cycle of rapid (re)development, followed by experimental manipulation and evaluation of the interface.

At the extremes, there are two approaches to getting the required flexibility. One is to code quick mock-up prototypes in a general-purpose programming language.

The second is to develop a special high-level language or "workbench" of tools for creating interface prototypes. The prototypes must operate in real time because they are to be tested in interaction with human subjects. But both approaches tend to run much slower than the ultimate system they mimic: the first, because the mock-up was developed for ease of programming rather than as an algorithm optimization research project; the second, because of the language overhead.

The most effective solution is to make more machine cycles available in the experimental stage than would be needed in production applications. Interface research relies on excessive computational power to provide the freedom to explore and evaluate alternative interface designs. In the future, these computational needs can be expected to grow.

Much of human-computer interaction research involves modeling the computer user in interaction with the computer. These models are of major importance to the theoretical development of the field of human-computer interaction. They also form the basis for new paradigms in intelligent interfaces, that is, interfaces that adapt themselves to the specific abilities and/or requirements of the individual user.

Computational power is particularly important in research on computer-based instruction. The machine must simulate both the student and the tutor and must be able to respond to inelegant user dialogue. Even the current powerful machines are capable of handling only small parts of such a system. (cf. in arithmetic, Burton & Brown, 1982; Brown, Burton, & DeKleer, 1982; in medical diagnosis, Clancey, 1982.)

Falling as they do within AI and cognitive science in general, these efforts in user modeling and intelligent interfaces have requirements similar to those already discussed: a need for interactive development environments and a powerful symbol manipulation machine. The major additional requirement, quite general to interaction research, is that actual interfaces must be tested with real humans, and therefore in real time. In typical AI efforts, it is sufferable to have a query answered in 30 sec. In testing an interface with a human user, response times of less than a second are necessary. Considerably more computational power is therefore necessary for an intelligent interface than for an "equally" intelligent non-interactive AI program.

Summary

Research in human-computer interfaces will need supercomputer resources. In general, there is a special need for fast prototyping of fast-running interfaces, requiring both clever software and high-speed hardware. Research in intelligent interfaces and AI modeling in human-computer interaction will require resources comparable to those outlined in the Cognitive Science section. Certain information retrieval contexts will need supercomputers, principally for statistical manipulation of very large

data bases. Tightly coupled human-computer interfaces will require high bandwidth input/output and the support of supercomputers.

CHALLENGES IN USING SUPERCOMPUTERS

A large portion of a researcher's time will be spent in theory and program development. This activity would include the repeated modification of programs and data, and thus would need greater front-end support, in the form of better editors, debuggers, etc. It would be essential to have moderate bandwidth (1200-9600 baud) real-time channels to good software on the supercomputer, or to some front-end smaller computer, tightly packet-switched to the supercomputer (local either to the supercomputer or to the user). Software development to make front-end computer use smooth and simple would be considerable, particularly in the context of debuggers.

Many applications in the areas of statistics and mathematical modeling would need only to ship programs and perhaps extensive data to the supercomputer and receive in return relatively small quantities of results and output. These applications need minimal facilities other than the supercomputer, some high-volume packet channel (e.g., tape) for shipping data and programs, and a low bandwidth interaction channel to direct execution. Some computer interaction research will require a real-time interactive environment (of various bandwidths) for running experiments on sophisticated interfaces.

Finally, it is imperative for cognitive science that new computer architectures be developed that are optimized for symbolic computation of the sort used in LISP or PROLOG.

NETWORKING

Optimal use of supercomputers requires that remote sites have substantial hardware and software facilities, including high bandwidth connections. One model for this would be a local node consisting of powerful workstations to handle all local editing, preliminary debugging, user interface, and graphics hardware. Software and high bandwidth communications provided for rapidly shipping data and programs to and from a remote supercomputer should be an integral part of any supercomputer usage. One possibility would be to interconnect the remote users over a network similar to current networks in which processes can be assigned to remote processors. In this case, a remote supercomputer would be assigned certain processes, and the results shipped back to interact with a local machine that would handle the simpler computational problems and user interface. Examples of this kind of interprocessor communication between their Crays and local workstations already exist at Los Alamos National Laboratory.

Although a bandwidth capable of effective file transfers will be important for a useful interaction of the nodes

with the central facilities, higher and lower bandwidths will be useful for certain applications. Some users will find that 1200-baud connections to remote machines will prove adequate. These usages will largely be those involving very large runs of relatively stable programs. Much work in psychology, however, will require substantial program development, which is poorly carried out at such low bandwidths. A small group of our users will require graphic output of a magnitude that will require exceptionally high bandwidth. This usage is probably extraordinary, and the cost for its development should probably be borne by the specific project and not be considered as part of the central facilities required by most users.

From the viewpoint of the laboratory perceptual scientist, the capability of generating and receiving graphic images over a short time period is essential. At present, the bit load of graphic images on existing communications networks is too large to permit the viewing of continuous motion. But, by storing graphic images computed at a central site, it will be possible to organize and then present at acceptable rates images that appear as a moving display.

For network transmission of graphic images, speed of transmission is essential because of the large amount of information that must be sent. For example, a small graphic image is defined by a bit map of 2 megabits ($512 \times 512 \times 8$ bits). At typical network transmission rates of 9600 baud, a single image requires nearly 4 min of transmission time. By comparison, a computation time of only 30 sec is needed to generate a single moderately complex image by the Digital Productions Cray-XMP.

A single image is but a small part of the motion sequences typically used in perceptual studies of motion. Often a stimulus sequence requires 100 frames of images presented at a rate of 40 frames/sec. For this stimulus sequence, the network transmission time is 400 min, or $6\frac{2}{3}$ h, far too long for effective communications. However, at satellite transmission speeds of 200K bits, this time can be reduced to only 20 min, a time quite acceptable to the laboratory scientist.

CONCLUSIONS AND RECOMMENDATIONS

Many of the research problems under study by psychologists are computationally intensive. Some of these problems are essentially numerical and for these the Class VI machine will offer substantial computational assistance. Another class of problems is best characterized as symbolic, and, although such problems are computationally intensive, the architecture of the Class VI machine is less than optimal. Supercomputers of a different architectural design seem to be required for these problems. Additionally, some computational problems faced by psychologists are largely parallel in nature and might be best served by a supercomputer with a parallel architecture.

We recommend supporting not only Class VI machines

but other architectures as well. It is imperative for cognitive science that new architectures optimized for symbolic processing be developed.

Additionally, the successful use of supercomputers will require the support of local workstations, high-capacity communication channels, and extensive software development. Additionally, some applications (i.e., graphics) will require local specialized hardware. We recommend that the NSF provide for local needs when furnishing supercomputer access.

REFERENCES

- ACKLEY, D. H., HINTON, G. E., & SEJNOWSKI, T. J. (1985). A learning algorithm for Boltzmann machines. *Cognitive Science*, 9, 147-169.
- ANDERSON, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- BROWN, J. S., BURTON, R. R., & DEKLEER, J. (1982). Pedagogical, natural language, and knowledge engineering techniques in SPOHIE I, II, and III. In D. Sleeman & J. S. Brown (Eds.), *Intelligent tutoring systems*. New York: Academic Press.
- BURTON, R. R., & BROWN, J. S. (1982). An investigation of computer coaching for formal learning activities. In D. Sleeman & J. S. Brown (Eds.), *Intelligent tutoring systems*. New York: Academic Press.
- CLANCEY, W. J. (1982). Tutoring rules for guiding a case method dialogue. In D. Sleeman & J. S. Brown (Eds.), *Intelligent tutoring systems*. New York: Academic Press.
- GREEN, B. (1957). The use of high speed digital computers in studies of form perception. In J. W. Wulfeck & J. H. Taylor (Eds.), *Form discrimination as related to military problems* (NRL Publications 561). Washington, DC: National Academy of Sciences.
- NATIONAL SCIENCE FOUNDATION. (1984). *Access to supercomputers*. Washington, DC: Office of Advanced Scientific Computing.
- NEWELL, A., & ESTES, W. (1983). Cognitive science and artificial intelligence. In *Research Briefings 1983*. Washington DC: National Academy Press.
- NEWELL, A., SHAW, J. C., & SIMON, H. A. (1958). The elements of a theory of human problem solving. *Psychology Review*, 65, 151-166.
- NEWELL, A., & SIMON, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- WRIGLEY, C. (1957). Electronic computers and psychological research. *American Psychologist*, 12, 501-508.

Appendix A Brief Introduction to Supercomputers

There are three features of currently available Class VI supercomputers such as the Cray 1S or the Cyber 205 that make the computing environment of these machines decisively different from previous computers: a different architecture (vector pipeline arithmetic), a larger main memory, and a faster execution rate. These new features have a profound impact on both the type of problems being solved and the programming techniques used for the development of software.

Faster cycle time and more main memory mean that larger problems can be solved in shorter time. Thus, supercomputers have made it feasible to solve more realistic models by computer simulation. Some spectacular quantitative improvements in the type of models now amenable to computer simulation have been made in areas such as aerodynamics, oil reservoir simulation, structural analysis, etc.

These trends will continue with the supercomputers that will reach the market within the next couple of years. The already announced models like Cray-2, Cray-3, and ETA Systems' 10GF machine will have even larger memory capacity. In addition to that, new supercomputers will have an extended memory capa-

bility such as that of Cray's solid-state storage device. Generally, we can expect a trend towards a hierarchical memory organization.

Execution rates will also increase. The increase will be due in small part to hardware improvements and faster cycle times, but most of the added speed will be achieved by a moderate amount of parallelism. The Cray X-MP (already on the market) comes with either two or four CPUs, and ETA Systems' 10GF machine will have eight CPUs. Multitasking will be available to the user; however, no supercomputers with massive parallelism are yet in sight.

Some of the advantages of supercomputers can be reaped by straight conversion of existing software. However, this will seldom result in all machine resources' being used optimally. The novel vector architecture complicates the picture, but it offers added improvement in speed. The vector architecture, therefore, has important implications for the user program. Algorithms must be selected and data must be organized to utilize the vector capability, and thus to realize the machine potential.

This is a new situation for the development of software. In the past, hardware considerations, algorithm design, and application programs could be considered as separate and independent activities. Now, they all overlap, and it is virtually impossible to use a supercomputer successfully while ignoring one of the three areas. This new situation creates a software dilemma for supercomputers.

One solution to this dilemma is the use of computational kernels in the development of software for supercomputers. A computational kernel is an optimized, high-performance implementation for a particular machine of a frequently occurring subtask, for example, a Fast Fourier Transform. For each new architecture, a new implementation of these kernels has to be provided. But with a standard set of kernels and calling sequences, only the kernels need to be changed when a given program is to be converted to a new architecture. With a set of kernels at hand, existing software can be converted to supercomputers with significant performance gains and with a minimum of user effort involved.

Most of the advances in the application of supercomputers have been made in areas of computationally intensive problems in the physical sciences and engineering. This is due largely to two facts: The software environment of Class VI supercomputers is mainly a FORTRAN environment, and the overall design of these computers is geared towards "number-crunching" applications. It would be wrong, however, to equate this somewhat restricted use of supercomputers today with a limited applicability of supercomputers to different problems. Recently, advances have been made to use supercomputers in artificial intelligence and graphics applications. The potential of radically new uses of the computing power of supercomputers to solve new problems is clearly visible but largely untapped.

Appendix B Access to Supercomputers

The document "Access to Supercomputers" (National Science Foundation, 1984), available from the NSF Office of Advanced Scientific Computing, describes the procedures for obtaining access to Class VI computers.

In April 1984, the NSF announced the formation of the Office of Advanced Scientific Computing (OASC). Under the direction of Dr. John W. D. Connolly, the office is divided into two program areas: (1) Supercomputing Centers, Program Direc-

tor, Dr. Lawrence Lee and (2) Networking, Program Director, Dr. Dennis Jennings. The office was formed to: (1) increase the access to advanced computing resources for the scientific and engineering research community, (2) promote cooperation and sharing of computational resources among all members of the community, and (3) develop an integrated research community with respect to computing resources.

On the basis of a documented need among universities for improved access to supercomputers, the Foundation initiated a program to facilitate access to Class VI computers for scientific and engineering researchers.

The NSF issued a Project Solicitation that encouraged academic institutions, nonprofit and profit organizations, and federally funded research and development centers to submit proposals to supply blocks of time on Class VI machines to NSF researchers. At the same time, the Foundation encouraged the submission of unsolicited proposals from individuals or groups of investigators requesting access to Class VI computer services.

If a researcher is submitting a proposal for initial NSF support of a research project that includes a need to access a supercomputer, the request for access should be an identifiable component of the overall proposal. If a researcher has an active NSF grant and needs to access a supercomputer to make further progress on this research project, then he or she should submit a request for a supplemental award.

Proposals requesting access to Class VI computer services will be received by the Foundation and processed essentially in the same manner as other NSF research proposals (see "Grants for Scientific and Engineering Research," NSF 83-57). All requests for new awards must be peer reviewed.

Because this is a new program, it is expected that the bulk of the initial awards for supercomputing time will be made in the form of supplements to existing NSF research projects. Over the longer term, however, researchers are encouraged to include requests for supercomputing time as components of their proposal.