Doing physics with computers

High-energy physics

Experimentalists in particle physics have long regarded computers as essential components of their apparatus and now theorists are finding that significant advances in some areas can only be accomplished in partnership with a machine.

Michael Creutz

Particle physicists have long been conspicuous consumers of computer time. Today's complex experiments generate prodigious quantities of raw data that require hours of cpu time to be reduced to a form comprehensible to human colleagues. As figure 1 shows, a multitude of computer cathode-ray screens often dominates the control room of a modern detector. The practicing high-energy experimentalist inevitably acquires an office cluttered with tapes and computer printouts, confering on the owner the image of a computer "hacker."

Michael Creutz is a physicist at the Brookhaven National Laboratory, Upton, New York.

The elementary-particle theorist, on the other hand, has traditionally projected a rather different image. Any problem requiring the aid of a machine for solution has been regarded as too earthy for a pristine intellect. Those theorists daring to sit before a terminal have risked the label "phenomenologists."

Although the experimentalist remains by far the primary user, the last few years have revealed an insidious infiltration of numerical techniques into even the more abstract sectors of the theoretical community. The results have in some instances been spectacular. This revolution has come about largely because computers have become, to use a catchword, "user friendly." Theoretical progress depends on numerous brief encounters with predominantly crazy ideas, out of which only a few will flourish. The theorists have only

Computer displays dominate the view inside the trailer housing the control room for E734, a large neutrino scattering experiment at the Brookhaven Alternating Gradient Synchrotron. (Photograph courtesy of E734 Group.) Figure 1.

recently realized that modern interactive computer systems allow toying with concepts on a time scale comparable to their attention span. Whereas the experimentalists have come to regard the computer as an essential component of the apparatus, the theorists are now becoming acquainted with a new tool of unmapped potential.

We shall see that to be of help to particle physicists, future generations of supercomputers must offer several orders of magnitude more power than current machines. Small personal processors will be able to take over most current-level calculation. On the other hand, special-purpose devices offer exciting prospects for particle physics which need to be pursued.

Needs of the experimentalist

Experimental elementary-particle physics requires numerical processing at several levels, the highest of which is the final offline reconstruction of events from raw experimental data. With a complex collection of detectors, the analysis of a single collision between incident particles can require as many as 10^7 machine instructions, which would take about one second of machine time to execute on a large mainframe such as a CDC 7600 or an IBM 3081. For low event rates, as in neutrino experiments or experiments using electron–positron colliders such as PETRA or PEP, this execution time presents no major problem. But this time becomes large when the experiment involves several million events. Many fixed-target hadronic experiments—such as at Fermilab, or in high-luminosity storage rings, such as the CERN Intersecting Storage Ring or the Colliding Beam Accelerator under construction at Brookhaven—may record of the order of ten interesting events per second. For these experiments the analysis of data from a few months of running can tie up the equivalent of several 7600s for more than a year. The associated demands are not only on cpu time; indeed, the data alone are voluminous. A single experiment might produce some 10^{12} bytes of raw data, requiring several thousand magnetic tapes for storage.

Actually, these demands are to some extent a consequence of what experimentalists perceive as the realistic limits of computational power. The fundamental interaction rate in hadronic experiments is normally considerably larger than what can be recorded for future analysis. For example, ISR has run with a collision rate of 5 million per second while the rate at CBA should increase to fifty million per second. Only after clever triggering on interesting classes of events can this rate be reduced to a manageable 10 HZ. Historically, the richest new physics has appeared in rare events. As just one example, the recent observation of the W meson at CERN relied on extracting five special events from 10^9 proton–antiproton collisions. Figure 2 shows a block diagram for the data-acquisition system in this experiment.

Triggering in and of itself represents another aspect of computation in particle physics. Input from various trigger counters must be immediately processed to determine whether or not to fire the entire apparatus and record an event. Such decisions are increasingly being aided by microprocessors that are an integral part of the experimental setup. These devices may be either an assortment of electronic gadgetry assembled by the experimentalist himself (see box), or a commercial product interfaced with the various detectors. If an online computer is available, it can be used for other useful tasks such as a preliminary analysis of the data. In this role the computer both reduces the final recorded volume of data and permits continuous monitoring of the system operation. Once the computer becomes a part of his apparatus, the

Specialized computers for high-energy experiments

Thomas Nash
Fermilab

In the near future, new experiments in high-luminosity fixed-target and colliding beams will require so much analysis by computers that commercially available systems may not suffice. This situation, which has been developing for some time, has already resulted in a number of efforts that have produced specialized processors with very impressive credentials. The increasingly widespread interest in this approach is demonstrated by the accelerating frequency of workshops and conferences on the subject of specialized processors: CERN, May 1981; Padua, Italy, March 1983; Brookhaven, May 1983; Fermilab, Fall 1983; Mexico City, January 1984. Intensive hardware development is underway at such laboratories as SLAC, Nevis (Columbia University), and CERN and SIN in Switzerland. Fermilab has formed the well-supported "Advanced Computer R&D Program" to focus efforts on specialized computing systems and to establish interactive relationships with supercomputer research in industry and university computer-science departments.

The increase in computing load needed for experimental analysis results from a number of factors that both permit and encourage the study of rare, complex multiparticle states in the presence of a background with a high data rate. Among these factors are the higher energy and luminosity available now from accelerators and the large solid angle and high-speed digitizing electronics of modern detectors. Most important is the interest of physicists in complex states arising from the decay of heavy fundamental particles, such as heavy quarks and intermediate vector bosons. The computer time needed to reconstruct events goes up very rapidly as the complexity increases, and because of the high data rates, there are orders of magnitude more events to process.

The special systems that have been built to date are already important in dealing with present computing loads for experiments. As an example of the magnitude of the requirements typical of present-day large experiments (though smaller than many planned for the next few years), I will consider some approximate numbers for data and computation rates for one active Fermilab experiment.

For each one-second beam spill, which occur every 10 seconds, there were some 10^6–10^7 beam–target interactions of one kind or another. Thus, in the 750 hours of recording data, there were a total of 10[illegible] interactions from which a series of "triggers" selected a total of more than 20×10^6 events. About 1500 digitized data words for each event were written on one of a thousand high-density tapes (carrying 6250 bits per inch). Typical of most modern experiments, this one used a multilevel triggering scheme. The circuits generally consist of large racks of interconnected modules, each performing specialized functions in parallel with dozens of other modules; the type of electronics they use is generally referred to by the industry standard (such as "NIM" or "CAMAC") that allows them to be interconnected. The first levels used "conventional" NIM fast logic to select, in 150 nsec, 5000 of the interactions per spill to be digitized and stored temporarily. Ten bins of commercial NIM modules, including simple coincidence, linear-summing and discriminator circuits, were involved at this level. Thirty crates of CAMAC electronics digitized thousands of channels of data from phototubes and wire chambers. A modular "data-driven" trigger processor (to be described below) acted as the highest-level trigger and carried out a sophisticated reconstruction algorithm on part of the data. The processor made the final selection of about 100 events to be recorded on tape by an online minicomputer.

The offline reconstruction of the raw digitized data into physics parameters (track momenta, angles and so on) is being carried out on Fermilab's three Control Data Cyber 175 computers as well as an IBM 3033 in Ottawa, six SLAC-designed 168E emulators (to be described below) at the University of Toronto, and three VAX 780s at U.C. Santa Barbara, the University of Colorado, and Carleton University in Ottawa. The dispersed panoply of computers that had to be brought to bear on these data is a symptom of the problem we are describing.

The total number of equivalent floating point computer operations in this example is staggering—over 10^{13} operations are required in the offline reconstruction. To this should be added the computation carried out by the high-level trigger processor on about a billion events—equivalent to 3×10^{14} operations before the offline computing began.

One of the earliest and most important specialized computers for high-energy physics was developed at SLAC, starting in the mid-1970s, initially for reconstructing data from SLAC's LASS spectrometer. This system, known as the 168E, is designed to emulate, bit for bit, the instruction set of the faster, much larger, and more expensive IBM 370-168. The emulation has the advantage that one can use the IBM 370-168, with all its special features, for debugging a FORTRAN program, which can then run on the 168E to process large volumes of data at perhaps ⅔ the speed, but for an order of magnitude less expense. The 168E, for example, uses smaller arithmetic units than the mainframe, which are microprogrammed to emulate the 370-168 instruction set. They are made of bit-slice microprocessors as well as smaller-scale integrated circuits; the bit-slice microprocessors (AMD 2901) work in parallel on small slices of long words, each in a separate circuit.

Because it is highly cost effective, and because its programs are easily debugged, the 168E has reached the status of a universal workhorse for high-energy physics experiments at SLAC, CERN, Fermilab and elsewhere. The popularity of emulators has led to a joint SLAC–CERN effort to develop the 3081E, a new and more powerful system keyed to the instruction set of the IBM 3081.

Some 168Es have also been used in online highest-level triggers. This demonstrates that there is only a soft distinction between online high-level triggering and offline reconstruction processing. Once the data rate has been reduced sufficiently, the recording medium (tape) can be inserted after any processor.

The M7, a processor important for its first use of a set of instructions specialized for track reconstruction, was built at Fermilab in 1977. Although intended primarily as a trigger processor, this system has, like the 168E, crossed the line and demonstrated its versatility in offline reconstruction. The M7 is a stored-program device that has selected features of a general-purpose computer. The M7 architecture and instruction set were based on a careful analysis of the essential computations involved in track reconstruction. In one major machine cycle (110 nsec) the M7 can carry out instructions of the form $E=A_i\times C\pm B_j\times D$ for any of a sequence of values A_i and B_j with constants C and D; a memory with four "ports" provides the operands for the instructions.

New kinds of computer "architecture"—that is, the internal operational structure of a computer—promise tremendous computational power. The traditional Von Neumann architecture involves operating on instructions following a sequence determined by a stored program. Each operation works on data from and puts answers into registers or memory before the next operation is started. A very powerful new architecture is best described as modular and data-driven: A large number of modules working in parallel process data (generally the results of operations from other modules) when all required data are available; their outputs in turn become available to other modules. Each module only waits for data specifically required at its input, minimizing process time. Operation is controlled by a system of "Ready" levels set by each module when its activity is complete. The first such processor was designed at Fermilab in 1978 and called the ECL-CAMAC Trigger Processor System. This system is in use on several experiments—the cover of this issue shows one example—and some modules of the system are now commercially available.

Programming such processors is done by interconnecting inputs and outputs of modules with multi-twisted pairs of wires or with transmission lines. Far more powerful than fast logic, these modules operate on multi-bit words, finding tracks, calibrating pulse heights and computing complicated functions. Single ECL-CAMAC module operations typically take 50 nsec. Although these systems operate in a high-speed data-driven mode, system algorithms may include hardware processing of nested loops, conditional branching and subroutines—operations normally associated with serial-computer processing.

The enormous increase over conventional mainframes in cost effectiveness for special calculations that such systems can attain was suggested in the experiment I described earlier. An ECL-CAMAC processor with a hardware cost of about $100 000 routinely carried out an algorithm in an average 7 microsec that took 40 millisec on a $1.5-million computer. A major drawback, however, is the difficulty and inflexibility in programming such processors and in verifying their operation.

A long-term design effort at Columbia University's Nevis Laboratory has led to the construction of a major new data-driven system. Taking a very systematic approach, the Nevis design incorporates many improvements that will go a long way towards easing some of the programming difficulties intrinsic to modular data-driven systems. A processor of this family is already in use on a Fermilab experiment, and there are plans to use a larger-scale processor on an experiment to be run in the near future at Brookhaven and Fermilab. The intention is to reconstruct completely a multiparticle experiment on line in less than 10 microsec. This will be a very impressive achievement.

As I noted earlier, future activities in this area are expanding. Fermilab's Advanced Computer Program is now studying approaches that combine the ease of use of modern 16- and 32-bit microprocessors, programmable in FORTRAN, with the power of novel architectures and special instruction sets.

The specialized computers we have described have important potential applications for computing-intensive problems outside high-energy physics. Historically, many of the earliest developments in computing and digital electronics were driven by particle and nuclear physics and conceived by physicists. After a long lapse it appears that experimentalists are, by necessity, getting into the business again, and we can hope for the same ultimate benefits to physics and society.

physicist is forced to function as computer scientist. He must develop both the hardware and software necessary for the proper interfacing, and the interface communication must be rapid enough to enable him to separate the interesting events from the flood of the mundane. Experimental progress is thus tied to the development and incorporation of state-of-the-art fast electronics into complex ensembles of detectors.

The computational needs of the experimentalist actually begin much earlier than in the running and analysis of an experiment. Detectors need to be designed and their performance characteristics determined before they are placed into the system. Simulation of existing or potential detectors requires detailed numerical analysis. An entering particle has a certain probability of firing a particular chamber, another probability of scattering when passing through some other component, and so forth. In a Monte Carlo simulation, many hypothetical particles are allowed to enter the detector and their progress is followed stochastically using a pseudorandom-number generator. The Monte Carlo approach is the only practical method for calibrating the complex systems in use today.

To design a detector, it is useful to have some predictions for the distinguishing features of potentially interesting events. This requirement leads to another area of computing in high-energy physics which lies between the experimental and theoretical. The generation of hypothetical sample events that are either typical or rare and interesting is valuable in formulating a detector configuration as well as comparing final experimental results with theoretical models. The Monte Carlo method is also used to generate these types of events. The particle-theory picture of quarks and gluons predicts that constituents of the incident particles may first undergo a hard-scattering process and then are "dressed" into ordinary hadrons via soft gluon bremsstrahlung and quark–antiquark production. Given a particular model, one uses stochastic methods to simulate the evolution of the resulting cascade.

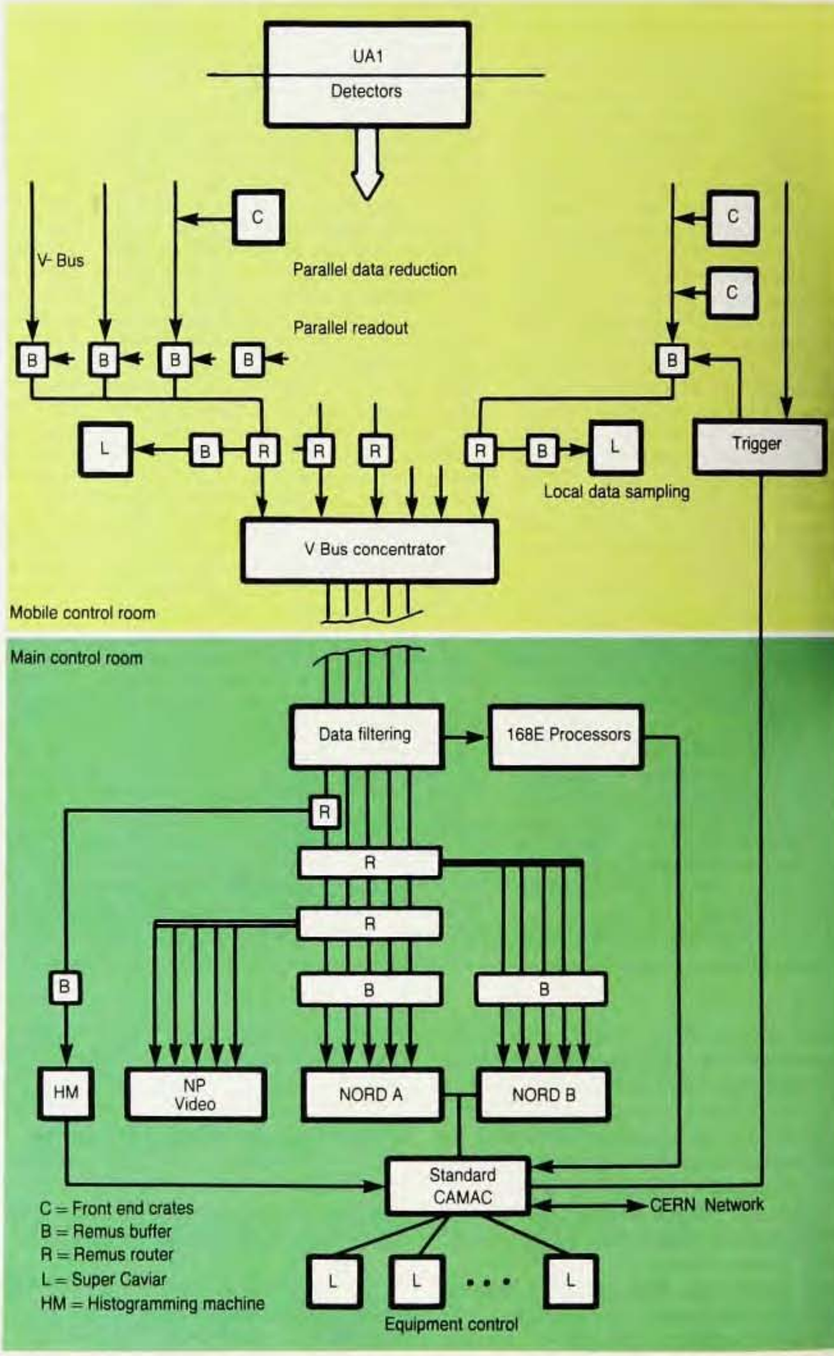
These predictions are important for two reasons that transcend the mere designing of equipment. First, one can test the various underlying assumptions of the model and perhaps thereby discover the underlying dynamics. This approach has been central to the analysis of hadronic jet production. Secondly, it provides us with predictions on cross sections in as yet unexplored kinematic regimes which in turn give us useful information on where to look for interesting phenomena. For example, from this kind of modeling we can get information on the production of new particles in future experiments, such as the heavy vector bosons that mediate the weak interactions.

Interest of the theorist

As we mentioned, computing has historically played only a background role in the development of theories of the elementary interactions. A breed of theorist known as "phenomenologist" takes a calculational scheme (which might once have been based on Regge theory but today more likely involves the currently popular quark–parton picture) and makes numerical predictions for direct comparison with measured cross sections. A more formal theorist might run across an integral that he cannot solve analytically or a differential equation that requires numerical solution. These are the traditional uses of computers by theorists, and they can often be avoided if they are not to one's taste.

About 15 years ago, theorists began to use computers in a new way. Theoretical calculations of complicated Feynman diagrams often involve extensive algebra with Dirac matrices. Such straightforward but tedious manipulations could in principle be done by machine; thus, algebraic manipulation programs such as REDUCE, SCHOONSCHIP, MACSYMA, and, more recently, SMP were developed. These codes represent a rather unconventional use of machines originally developed for data processing. Instead of numbers, one manipulates strings of characters and provides a set of rules for algebraic substitutions. The programs may even include tables of common integrals to remove still more drudgery from the calculation. In this way, calculations of increasing complexity have become possible; indeed, such programs have become essential tools for higher-order perturbative calculations in quantum electrodynamics or non-Abelian gauge theories.

In only the last few years a second new use of computers has surfaced which has resulted in a massive increase in the demands of theorists for cpu time. Monte Carlo simulation has become a powerful tool for studying the solutions to interacting quantum field theories. These calculations have concentrated on the gauge theory of the strong interactions (based on quarks interacting by exchanging gluons in a Yang–Mills gauge theory). The main result is that we now have rather compelling numerical evidence that this theory can simultaneously give rise to the phenomena of quark confinement (wherein an isolated free quark would have infinite energy) and asymptotic freedom (wherein the interactions between quarks become weaker as they move



Data-acquisition system of UA1, a large detector at the CERN SPS Collider. A set of Caviar and Super Caviar microcomputers control and monitor individual components. The Nord 100/500 computers form the general online computing support. The system also contains several 168E processers, emulators of the IBM 168. (Courtesy of S. Cittolin, UA1 Group.) Figure 2.

closer together).

These Monte Carlo simulation techniques have been borrowed directly from the solid-state physicist. Indeed, the particle-theory applications have served to emphasize the deep mathematical connections between these subfields of physics. The ties are revealed most clearly via Feynman's path-integral formulation of quantum mechanics. From a mathematical point of view, a path integral is identical to the partition function of statistical mechanics; that is, quantum field theory in four space–time dimensions is equivalent to the classical statistical mechanics of a hypothetical thermal system with four spatial dimensions. (In this analogy, the coupling constant of the field theory plays the role of the reciprocal of the temperature.)

The condensed-matter theorist has long used the Monte Carlo method to simulate classical statistical systems with an arbitrarily chosen Hamiltonian. The computer memory contains the numerical values for the degrees of freedom—for example, the directions of the spins in a model of magnetism. A random-number generator, weighted by the Boltzmann factor, then induces changes that mimic thermal evolution and fluctuation. By isolating various features of the dynamics, the physicist gains insight into phenomena such as phase transitions.

Particle theorists use the Monte Carlo technique as a numerical method for evaluating Feynman path integrals. In this way they can study the solution of their favorite field theory. Before proceeding to this goal, however, they must face the difficulty that quantum field theories are notoriously replete with ultraviolet divergences. If an integral is not mathematically well defined, a sure way to find out is to try to evaluate it numerically! Of course, infinities are an old subject in particle physics, and numerous regulators and renormalization schemes have been developed. However, most of these procedures are based on perturbation theory, the backbone of the spectacular successes with quantum electrodynamics. One starts calculating a scattering process with Feynman diagrams, and, upon finding an infinity, cuts off the corresponding integrals. Numerical simulation, on the other hand, is independent of perturbation theory. Consequently, most of the traditional regulators, such as those first used by Wolfgang Pauli and Felix Villars or dimensional continuation, are not directly applicable. For this reason all numerical work to date has used the "lattice cutoff" introduced by Kenneth Wilson of Cornell University. In this scheme, integrals over all space–time points are replaced by sums over a lattice of discrete space–time points.

The use of a space–time lattice as a regulator has several advantages:

▶ It serves as an ultraviolet cutoff. The continuum of space–time is replaced with a set of points separated by minimum length a, which in turn gives a maximum momentum and energy on the order of h/a. The world lines of quarks are then approximated by sequences of steps between nearest neighbor lattice sites, as illustrated in figure 3. Of course, this is just a mathematical artifice; at the end of any calculation we must extrapolate to the continuum limit of vanishing lattice spacing.

▶ The numerical treatment of lattice variables is straightforward. The fields simply become arrays in the computer memory.

▶ On a lattice the analogies with statistical mechanics become even more striking. We work with a four-dimensional crystal and the gauge variables have a superficial resemblance to spins in prototypical models of magnetism.

The nonperturbative nature of the Monte Carlo method is actually essential for the quark confinement problem. Straightforward renormalization-group arguments show that the confining part of the interquark potential cannot be expanded as a power series in the coupling constant: It displays an essential singularity, vanishing faster than any power as the coupling is reduced to zero. It is this non-perturbative nature of confinement that makes the lattice regulator so important, even beyond the context of the numerical work.

The results of these simulations have been spectacular. Not only do we have compelling evidence for quark confinement in the non-Abelian gauge theory of the strong interactions, but we also have been able to calculate several observables characterizing the solution of the theory. The first of these is the relative strengths of the short- and long-range interquark potentials. The calculated values are in reasonable agreement with phenomenological values from hadronic spectroscopy and inelastic scattering experiments. Further calculations provide evidence for a rich spectrum of bound gluon states (glueballs) with masses in the 1–2 GeV range. There appears to be a deconfining phase transition at a physical temperature of approximately 200 MeV, a value potentially attainable in high-energy heavy-ion collisions. These are all non-perturbative results, which only a few years ago would have been regarded as hopelessly difficult to calculate even qualitatively.

The most secure of the Monte Carlo results concern the quantum dynamics of the gluon fields and treat the quarks only as static sources. Calculations including the full relativistic quantum dynamics of the quarks represent an intensely active area of current research; indeed, there are severe obstacles to be overcome. It is not yet clear whether the difficulties with quarks are primarily due to inadequate computer technology or to our theoretical ignorance in formulating the problem. We do know that given sufficient computing power, one can calculate with the quantized quark fields. However, as currently done, the calculation requires the repeated evaluation of determinants of extremely large matrices. A few daring theorists are proceeding with this sledgehammer approach. To make the calculation feasible, they have devised various tricks using the fact that the relevant matrices contain many zero elements and at any stage of the calculation only a few of the remaining elements are modified. The results of these efforts are eagerly anticipated.

An interesting simplification has recently provided several exciting results on hadronic spectroscopy. This "valence" or "quenched" approximation introduces a small number of dynamical quarks into the lattice and allows them to propagate including all interactions with the gauge fields. The approximation consists of neglecting the possibility that further quark–antiquark pairs can be excited from the vacuum. The only quarks around are those introduced initially. These calculations give hadronic masses that are qualitatively similar to those observed physically. In particular, the pion is considerably lighter than the remaining hadrons. The result, theoretically related to what is referred to as chiral symmetry, has been notoriously elusive in simple quark models. Recent careful studies indicate that the approximation has some shortcomings—the nucleon, for example, comes out a few hundred MeV too heavy. Nevertheless, it is remarkable that such a crude approximation works so well. This may not be a new puzzle; indeed, it is not clear why any simple valence quark model should work at all.

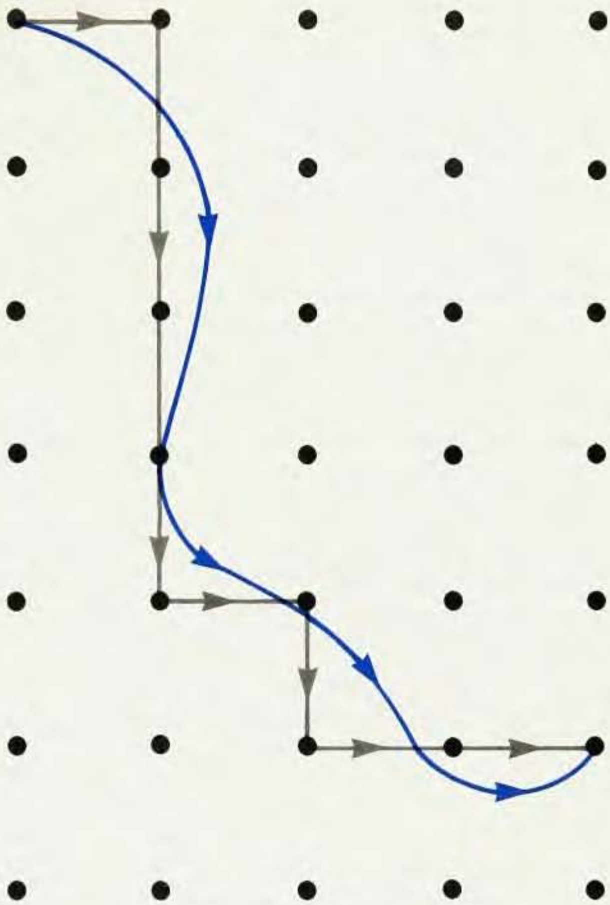
Future directions

Computer technology is escalating at an astounding rate. What speculations can we make about how this will affect the particle physicist? Undoubtedly we have not even yet dreamed of the most important uses in the future. Ten years ago no one imagined the potential of Monte Carlo methods for field theory. Of course, in addition to speculating we must continue to make definite plans for the future, as our calculations continue to push current technology and future accelerators promise to place even more challenging demands on data processing.

There are several directions that particle physicists can take to meet future computational needs. Dramatic cost reductions in hardware are making respectable computing engines available to the individual scientist. Microcomputers now provide the power of a respectable fraction of a moderate-size mainframe, but without the headaches of unpredictable queues or system crashes. Personal computers will presumably function as sophisticated calculators, doing many of the mundane numerical chores of both theorists and experimentalists. However, as physicists learn more of what can be done, they will inevitably be enticed into calculations too large for reasonable turnaround on a personal machine.

Going up somewhat in scale, respectable intermediate systems, such as the VAX 11-780, are now in the cost range accessible to small groups of physicists. As with personal computers, major advantages of a group owning its own machine are enhanced reliability and elimination of the bureaucratic haze that usually surrounds large computer centers. With a long program to run, one just leaves the computer on when one goes home in the evening. Indeed, many of the longer Monte Carlo studies in lattice-gauge theory have been done by locating idle hours on VAXs. Even though the calculations could have been done an order of magnitude faster on a CDC 7600, the cost would have been prohibitive.

We note that large mainframes offering only a single order of magnitude of computational speed or capacity over small machines are of diminishing value to the physicist. A changing role will nevertheless ensure the survival of the computer center, which will serve primarily as a network node for software and information management. The numerous local computers of a research establishment will rely on this service for access to the latest editions of general program libraries. For basic computation, however, larger computers must grow several orders of magnitude in capacity. The primary uses for such machines will be for massive simulations or other large calculations that require reasonable turnaround times. Although electronics alone are getting faster, the most promising hope for the needed vast increases in power involves developing systems with massive parallelism. Keeping many interlinked microprocessors efficiently occupied is a non-trivial issue for computer architecture as well as software and is currently the focus of active research in computer science (see box).



Quark world-lines are approximated by a sequence of steps between neighboring sites in lattice gauge theory. The results are then extrapolated to the continuum limit of vanishing lattice spacing. Figure 3.

For problems requiring major number crunching with little input and output, array processors may provide a low-cost interim solution. These machines incorporate efficient "pipelining" of instructions to keep the various processing elements as busy as possible. They are attached to a conventional mainframe computer, through which all input and output is channeled. Although a few of these machines are in use by theorists, most notably at Cornell University, the software to make them easily exploited is still under development.

The rapid drop in cost for electronic components has opened another avenue to cheap computer cycles. A few thousand dollars worth of parts can provide a circuit that will "emulate" the processor of a standard mainframe (see box). That is, for a given set of machine instructions, the emulator will execute them with the same output, bit by bit, as the machine it imitates, although it may do so more slowly. Such a gadget is useful when one wishes to apply the same code to a large number of "events." One first uses a well-optimized compiler on the corresponding mainframe to produce an efficient machine code. This is then fed to a set of several emulators, each of which goes to work on a separate event. In experimental work these events would be the raw output of the apparatus while in theory each could represent a different value of coupling constant. Several current experiments are productively incorporating such emulators.

As electronic components have become cheaper and more sophisticated, it has become feasible to build machines specially designed to do a single calculation. By not asking for general applicability, one can design a machine that is both simpler and faster than the large general-purpose machine. Such devices can be coupled with a conventional computer, which thus acquires a new component for special tasks. For example, theorists at Columbia University have built a multiplier for SU(3) matrices. When coupled with a small conventional computer, the system performs with the speed of a much larger machine for Monte Carlo calculations. Carrying the special-purpose processor idea to the extreme, we have the concept, currently being developed, of a "silicon compiler," that is, a computer that designs a large-scale-integrated circuit chip whose sole function is solving a given type of problem.

Of these possible directions, which holds the most promise for the particle physicist? The answer is, of course, neither known nor likely to lie along any single path. There are, however, some clearly discernible differences between the experimental and theoretical needs. When reducing data from an experimental event, the computer encounters extensive conditional branching due to the complex ways the detector can fire. In addition, there are substantial input and output demands when several million events must be fed into the computer and correlated with each other. In general these requirements render it difficult or impossible to make efficient use of architectures incorporating sophisticated vector parallelism. In particular, it is not clear what the new generations of supercomputers will be able to offer for this type of data processing, which requires multiple instruction and data streams. Hence, experimentalists have given extensive emphasis to the "unsophisticated" approach of using many independent emulators. Similarly, the input–output limitations reduce the attractiveness of array processors in this area. However, experimentalists are more enthusiastic about the possibilities of fully parallel supercomputers, such as New York University's Ultracomputer and the HEP from Denelcor.

The simulations of interest to the theorist, on the other hand, are readily implemented on multiprocessing systems. As with other simulations of classical and statistical systems, the evolution of any particular degree of freedom in a large array depends directly only on its neighbors. This means that several processors can concurrently do the same calculations on distinct parts of the system. A group of theorists at the University of Edinburgh are thus productively doing lattice calculations on the Distributed Array Processor, a device containing a total of 4096 interlinked microprocessors (see the article by Jorge Hirsch and Douglas Scalapino, page 44).

Unfortunately, like particle accelerators, state-of-the-art supercomputers will inevitably have access priorities. Accordingly, their attractiveness for casual use will be reduced, effectively limiting these machines to calculations that cannot be carried out with reasonable turnaround time on more personal machines. Unless a conceptual breakthrough appears, Monte Carlo simulations with dynamical quarks will necessarily be done on these machines; for other purposes, however, the mainframe will decrease in importance. The new generation of small computers will take over most of the mundane theoretical chores such as algebraic manipulation.

For heavier loads, the array processors continue to tantalize theorists with their inexpensive computations. It will be interesting to watch the course of events as theorists gain experience on these devices and as software continues to develop. A few theoretical calculations have been performed on emulators developed by experimental groups. The interesting observation arising from these trials is the enthusiasm and psychological advantage experienced when the theorist has access to large amounts of free computing. This freedom encourages thought in new directions on computational problems that would not be considered via conventional channels.

Special-purpose machines for particle-theory calculations are under development at the Institute for Theoretical Physics at the University of California in Santa Barbara and in the theory group at Columbia University. A somewhat more general machine incorporating multiprocessing is under development by a collaboration of particle theorists with computer scientists at Caltech. The potential power of all these devices is exciting, but one faces the danger that by the time the machine is finished, theoretical interests will have changed. The fickle attitude of the theorist can be partially compensated by flexibility in the machine. As technology decreases the development time for new devices, this pioneering work at a handful of laboratories will evolve into widespread use of new hardware and software.

The role of computing in particle theory is so new and rapidly changing that future directions remain extremely vague. Most results so far have been obtained in a quick and dirty fashion on conventional mainframes. Rapid technological developments ensure that the importance of numerical work in theory will not diminish; indeed, in a few years the computer may well be as essential a tool to the theorist as it now is to the experimentalist. As the unforeseen possibilities of massive computing power open up, we may well marvel at how the mere manipulation of strings of bits can have such a profound effect on our search for the ultimate constituents of matter.

* * *

Work described in this article was performed under contract number DE-AC02-76CH00016 with the US Department of Energy. □