Computational Chemistry: Getting More from a Minicomputer

Minicomputers are providing increasingly more computational power for the dollar as microelectronics technology advances. Thus, it was probably only a matter of time until researchers elevated the mini from its customary role of controlling experiments and interfacing them to larger computers located elsewhere to that of a full-fledged machine for scientific computation. Researchers at the University of California's Berkeley and Irvine campuses have, in fact, recently made this jump and are using minis to calculate the electronic structure of molecules and study the dynamics of molecular collisions.

While most observers are stopping short of predicting a revolution in the way scientists do their computing in the future, the trend toward the use of ever larger computers for such computation could reverse if the idea of using a minicomputer catches on.

Henry Schaefer and William Miller at Berkeley have, for the last 2½ years, been using a minicomputer for much of their theoretical work. Schaefer calculates the electronic wave functions of molecules and, using these, the potential energy functions for molecular interactions. Miller uses semiclassical (that is, not completely quantum mechanical) methods to obtain the trajectories of atoms and molecules that collide and react chemically. Both of these activities are in the class of what has been called large-scale chemical computation, which has been thought to require the largest available computers to accomplish.

Yet Schaefer and Miller find that their minicomputer not only can handle these computations but, depending on the particular calculation, can perform them for from one-eighth to one-third the cost of the same projects on the Lawrence Berkeley Laboratory (LBL) computer. While the minicomputer takes about 25 times longer to do the job than LBL's Control Data Corporation CDC 7600, the researchers are now able to do calculations that were previously prohibitively expensive, thus extending the range of problems accessible to them.

Donald Bunker at Irvine has a less expensive and slower minicomputer than that of his Berkeley colleagues. He and his associates have been using it for more than a year to do wholly classical trajectory calculations. Bunker reports that, for about equal financial investments, they have a ten times greater power to attack new problems with the mini than they have at the Irvine computer center. And, stimulated by the success of Schaefer and Miller, chemistry researchers at the University of Wisconsin, Madison, and at Bell Laboratories, Murray Hill, New Jersey, have completed studies showing that they can do large-scale chemical computation with minis for from one-tenth to one-half the cost of using their respective computer centers and are now in the process of acquiring minis for their research.

Along with advancing minicomputer technology, economics is inextricably bound up with the question of whether to go to a dedicated minicomputer or stay with a central maxicomputer. Three years ago, Schaefer and Miller looked at their computer budgets (about $25,000 a year each), at the cost of computing at LBL [about $1500 per hour when both central processor unit (CPU) computing time and input-output device time were accounted for], and at how much computing they thought they needed for their planned projects.

A Mini Saves the Day

Not liking what they saw, the two investigators began to seek alternatives, which led them to consideration of minicomputers. Close examination of the price and performance of the then available minis revealed that there were about a half-dozen candidates that might meet the computational requirements. In the end, Schaefer and Miller selected what is now the Harris SLASH 4 as qualitatively superior and submitted a proposal, which the National Science Foundation (NSF) accepted.

The system they proposed comprised a CPU with 64,000 words of 24 bits each of main memory (or 32,000 double precision words of 48 bits each for their most accurate computation), a 19-million-word disk to serve as auxiliary memory and hold the operating system (programs that run the computer), and various related equipment. The total price was about $130,000. Additional expenses included a service contract costing $1155 per month and miscellaneous items estimated at $300 per month.

Assuming the minicomputer was 1/64 as fast as the LBL computer, depreciating the cost of the system over a 4-year period, and running it about 20 hours per day, the Berkeley chemists had estimated the hourly cost of computation equivalent to that obtainable from LBL to be $443. In practice, the speed of their mini has been, on the average, 25 to 30 times slower than that of the central computer, so that the actual cost of equivalent computer time has been closer to $200 per hour.

Although there initially were several problems in getting the system to run because of defects in the hardware, recently the mini has been inoperable only about 1.5 days per month, usually because of problems with electromechanical devices, such as the teletype or the line printer purchased as part of the system. But, since there are no computer center personnel to help them, each researcher has had to become a computer operator as well as programmer. Moreover, one person has also had to become very familiar with the working of the computer in order to keep it operating at high efficiency; it is definitely not a black-box type of operation. Nonetheless, in the time they have had their minicomputer, Schaefer and Miller, together with their students, have produced about 40 research reports for publication, 33 of which used results from no computer other than the mini.

At Irvine, Bunker purchased a Hewlett-Packard 2108 for $27,000. Since his classical trajectory calculations do not need a large memory, Bunker did not buy a mass storage device, such as a disk. Bunker and his associates have already completed one major study that took about 4 months of nighttime running on their mini system. The same project, at the Irvine computer center, would have cost between $7000 and $8000, so they think that a substantial portion of the capital investment for a mini has been amortized on this one project alone.

The chemistry department at the University of Wisconsin currently pays an average of $200 per hour for computer time on the university's Univac 1110. Since the Univac is several times slower than the LBL computer, for example, this rate is not as inexpensive as it might first appear. Depending on which of several benchmark or test programs they ran, chemists at Wisconsin concluded that a newer and faster Harris minicomputer, the SLASH 7, would be 4 to 12 times less expensive to use than their computer center, thus increasing the amount of computation they can do by the same factor.

Rather than being dedicated to theoretical computation only, the minicomputer system, priced at about $150,000, will serve the group in three ways, according to Wisconsin chemists John Schrag and Philip Certain. Besides carrying out large-scale computations for the quantum chemists and x-ray crystallographers in the department, the mini can operate as a terminal for the university computer and it can do data manipulation operations, such as curve fitting.

The SLASH 7 also has a feature introduced some years ago by Burroughs for its large machines, a virtual memory. This feature means that 256,000 words of memory located on the disk appears to the computer as if it were in the main memory. When the machine is operated in this mode, provided that the virtual memory is not too large, it takes only 15

Chemists' First Try at Big Science

If high energy physics is the epitome of big science with multimillion dollar budgets, chemistry remains, according to some, a cottage industry. In their first major attempt to benefit from the pooling of resources in a cooperative effort, chemists are now inching toward the establishment of a national center for computational chemistry. Termed the National Resource for Computation in Chemistry (NRCC), the center is designed to serve the needs of all chemists whose computational needs cannot adequately be met by local computer centers. Operation of NRCC could begin by the fall of next year.

Jacob Bigeleisen of the University of Rochester, who has been chairman of a National Research Council planning committee for NRCC, sees two major groups of users for the center. The first comprises those chemists requiring large amounts of running time on the largest and fastest computers available. At present, for example, no university possesses computers of this type.

The second group is made up of chemists who could profitably use well-documented computer programs in their research but who, for lack of expertise or resources, cannot develop the programs themselves. Indiana University's Quantum Chemistry Program Exchange already serves as a repository for and distributor of such programs but does not now engage in developing them or putting them in any kind of standardized format.

Observers also point out that NRCC, which will gather in one place research experts in both computer science and in chemistry, should stimulate the optimal utilization of computers in chemistry and greatly enlarge the range of problems chemists can address.

The NRCC idea has been around for several years, having been first articulated by Harrison Shull of Indiana University and later brought to national attention by a National Academy of Sciences–National Research Council committee on computers in chemistry. It seemed to pick up more steam when Energy Research and Development Administration (ERDA) and National Science Foundation (NSF) officials became interested. Shortly thereafter, ERDA and NSF jointly requested the National Research Council to follow up on the earlier study by providing detailed recommendations for the organization and operation of NRCC. The National Research Council planning committee, which issued its report a year ago, recommended that NRCC be operated by a professional, in-house staff responsible to a director and a governing policy board made up of outstanding computational chemists. Selection of projects undertaken at NRCC would be by a program review committee. A users' committee consisting of professional staff and visiting scientists would also provide feedback on users' needs.

Since the NRCC policy board has not yet been selected, the planning committee recently sponsored a series of workshops to identify the types of projects that NRCC could best service and the facilities (hardware and software) needed to provide this. The recommendations of the workshops, when available this fall, will be purely advisory and nonbinding, however.

The planning committee also recommended that NRCC proceed in two stages. In the first stage of operation, no hardware is to be purchased. Instead, the emphasis will be on organizing the center and demonstrating its usefulness. Scheduled to be evaluated after 3 years, NRCC activities will probably be carried out at one of the ERDA national laboratories, many of which have the most up-to-date computational facilities and sufficient unused computer time to serve NRCC's requirements.

As now conceived, both NSF and ERDA will jointly support the center for the trial period. The National Science Board has already approved NSF participation in the project, provided that a suitable proposal is submitted. The two organizations are, in fact, in the final stages of negotiating a memorandum of understanding that will set forth the terms of their partnership. After the memorandum is agreed upon, proposals will be sought from the ERDA laboratories and reviewed by a committee of NSF and ERDA personnel. A site will then be selected.

Although the budget sent to Congress for fiscal 1977 by the Administration contains funds to support the first year of NRCC operation, it is not likely that activities could begin until the start of fiscal 1978, according to Richard Kandel at ERDA, even if the full amount is approved. The exact amount of funding will depend on the proposal accepted. The planning committee had projected that, from a first-year budget of $1.3 million, support would grow to about $2.4 million (in 1974 dollars) by the third year of the trial.

Depending on the progress of NRCC during the initial 3 years, observers see a number of possibilities for the second phase. The project could be discontinued, could be continued at the stage one site, could join a computer network, or could purchase its own computer facilities. In the latter event, ERDA would probably withdraw from the project, but officials view the prospect of purchasing hardware as unlikely. If the project continued at an ERDA laboratory, NRCC would not become subservient to the laboratory, but would remain scientifically autonomous. This point was strongly emphasized in the planning committee's report.

Although officials at NSF and ERDA and those on the planning committee are encouraged, there is, as yet, no firm commitment by any agency to proceed with the establishment of NRCC. No specific proposals have been received for anyone to rule on, although these are expected by this fall. Thus, for the moment, chemists' first venture into big science, while getting close, is not yet a reality.—A.L.R.

to 20 percent longer to run the benchmark programs than when only the main memory is used. This speed is within a factor of two of that of the central computer.

The Wisconsin group will probably hire a scientist to oversee the operation of the machine, and most users will have to worry only about their programs. But no consulting or other assistance of the type obtainable at a computer center will be available.

Computational chemists at Bell Laboratories are simulating the motion of the individual atoms in a molecule as the atoms relax toward their equilibrium configuration from various initial positions. This type of study requires several hours of continuous computer time, and the operator must be able to interact with the program as it proceeds. At Bell Labs, however, such long-running computation is usually done overnight or deferred for days until an open time slot is available.

According to John Tully at Bell, a high-performance minicomputer, costing between $180,000 and $190,000, could be constantly available to ten or so theoretical chemists there for about half the equivalent cost of using the laboratory's computer center. Moreover, such a minicomputer is faster than the Honeywell computer the center uses! A computer operator would be hired to operate the machine, as at Wisconsin. Tully attributes the somewhat lower savings at Bell Labs, as compared to the universities mentioned, to the absence of graduate students (and their attendant low salaries) at the former location. In fact, graduate student labor is not usually figured in as one of the costs of using a minicomputer.

Sources of the higher cost of using any computer center include the salaries of its professional staff, the expensive peripheral devices the center must maintain but which chemists use infrequently, and the effective subsidizing of certain groups of users, such as computer students, who do not pay their way.

Heretofore, theoreticians have believed that anything more involved than data manipulation required the use of the largest available computers—witness the drive toward the establishment of a National Resource for Computation in Chemistry to, in part, provide a place for chemists to do computation that is not possible elsewhere (see box). Campus and laboratory computer centers have also had a certain stake in encouraging this attitude, since a significant portion of their revenue is derived from computational scientists. A large swing toward the use of dedicated minis for large-scale computation could thus have a significant impact on the distribution of computing facilities and in the most extreme case might threaten the existence of central computer centers.

Fears of this type already have played a role in stimulating opposition to the use of minicomputers for large-scale computation. According to Charles Stevenson, who is now executive secretary of the statewide University of California Computer Policy Board (which oversees the purchase and use of computers), complaints from the Berkeley computer center that they would lose a major source of revenue were in part responsible for the initial refusal to permit Schaefer and Miller to try their experiment, even after the NSF had expressed very strong interest in the project. It was only after the university president, Charles Hitch, was persuaded to personally intervene that the researchers were able to go ahead. Ironically, Stevenson says the current trend of computer economics may soon force more computational researchers into using minis.

Opposition to Minis

At the University of Wisconsin, where the computer committee claims to seek a balance between centralized facilities and dedicated computers, chemists feel that their project met much resistance simply because they openly intended it to include large-scale computation. Interestingly enough, there are more than a dozen high-performance minis on the Madison campus. Although the purchase of each had been justified on the basis of its being used for controlling experiments or computer instruction, several are now used mainly for computation.

Some scientists think that fears for the future of central computers may not be altogether unfounded. For example, while the largest available computers may be cost effective under certain circumstances, says Schaefer, any computational chemist using a smaller, less efficient computer at a computer center is not getting his money's worth and should rebel by buying a dedicated mini. There are now several computers available that are comparable to the one at Berkeley. This will be even more true in the future because minis that are four times faster than Schaefer and Miller's machine are likely to be available soon for about the same price.

Peter Lykos of the Illinois Institute of Technology in Chicago, foreseeing a trend toward dedicated computer systems of several kinds, feels that advancing computer technology and the increasing importance of networks of geographically separated computers which can share their resources mean that no central computer facility could offer the computational chemist the services he could obtain more economically on his own. Lykos is now organizing a symposium on minicomputers and large-scale computation in chemistry to be held next year by the American Chemical Society.

Exactly how widespread the use of dedicated minicomputers for large-scale computation has become is unknown. While Tully, at Bell Labs, sees a dozen or more computational chemistry groups that could profit by owning their own minicomputer, but not a hundred or a thousand, the Harris Corporation claims to have already sold more than 100 high-performance minis to computational scientists in several disciplines. Most observers would probably agree with James Brown of the University of Texas, who says that these experiments have demonstrated the existence of a significant class of problems which can be solved more efficiently and cheaply with a minicomputer. But the position of the dividing line that separates these problems from others for which the use of minis is unwarranted depends on economic considerations which are subject to considerable change.

Don Secrest of the University of Illinois points out that, to be economical, the minicomputer must run close to 24 hours a day for up to 4 years, yet this level of performance has yet to be demonstrated. In addition, adds Secrest, the economics of computer centers could change in the next few years. One of the simplest such changes would be for computer centers to adopt what Secrest calls "honest billing," whereby the computational scientist pays only for the relatively cheap CPU time that constitutes the bulk of his demand on the computer, rather than subsidizing others who use little CPU time but depend heavily on expensive peripheral devices.

Most observers seem to agree that, because central computer facilities are needed to provide services to relatively unsophisticated users and because they offer a large array of devices and services that the owner of a mini could not obtain for himself, central facilities are not about to disappear. Thus, for now, the importance of the concept of using a minicomputer for large-scale computation seems to be that, for some scientists, it opens the way to calculations they could not have afforded in the past. For these people, the mini is the difference between doing and not doing a computation.—ARTHUR L. ROBINSON