# Contents

# 1

# The 'E' in ICCCRE

PETER LYKOS
Illinois Institute of Technology
Chicago/IL 60616 (US)

## 1.1 The 'E' in ICCCRE

The first ICCCRE was held at Northern Illinois University (DeKalb, IL) in 1971. Part of the funding came from the US National Science Foundation's *Education* Directorate. One of the main purposes of that Conference was to bring together those at the forefront of bringing computers to the practise of chemistry with those preparing the students of the day to become the professional chemists of the future.

As was reinforced by the ARTHUR D. LITTLE report (on the occasion of the 100th anniversary of the American Chemical Society in 1976) the ACS has two principal purposes:

1. the advancement of the science of chemistry
2. the education of chemists.

The ACS plays a role which is distinctive in the spectrum of major scientific disciplines. The ACS is the only major professional society in the USA supporting its members active in all the aspects of the practice of the science including industry, research and development, and education. The coupling and interaction between the several components of the chemistry establishment is facilitated—and enhanced—through the forum which is the ACS.

Enter the computer—that entry first publicly identified, in 1951, by the International Conference on Quantum Mechanical Methods in Valence Theory at Shelter Island, New York. Just eight years later, in 1959, there happened the recorded incorporation of a computer application as a regular part of a required undergraduate chemistry course (programming in machine language to do data analysis as part of a physical chemistry laboratory).

Computer-based enchancements to the practice of chemistry are being introduced and refined at an ever increasing rate. However, it is likely true that even those scholars leading the chemistry establishments toward greater and greater insights into the ever expanding discipline we call chemistry—even they, on the average—lag substantially behind in use of state-of-the-art technologies of computer science and engineering and in *realization of the potential for technology transfer* from discipline to discipline via the vehicle of computer hardware/software.

The problem (or loss of opportunity) is exacerbated in chemistry education in the USA for three reasons:

1. Universities are distinctive in the major emphasis placed on research activities which is large compared with the emphasis on improving the undergraduate program in chemistry. Those *most* likely to be at *the cutting edge* are *least* likely to be concerned with maintaining the quality—and the currency—of undergraduate chemistry.

2. Colleges, generally not having graduate programs, place greater emphasis on undergraduate education and, of course, the undergraduate research experience has been—and continues to be—the most valuable—the *integrating*—component in that process. But there the so-called computer-based enchancements brought into the educational process tend to be *computerization* of B.C. (Before Computer) problem-solving methods—hardly an attraction to bright young minds to enter a career in chemistry.

3. The quality of science and mathematics education and the level of communication skills of graduates of the secondary—or feeder—schools have fallen to an unacceptably low level—so much so that colleges and universities have had to give up some of the already too little time in the undergraduate chemistry curriculum for remediation programs.

The ACS has a joint Board-Council Committee on Professional Training (CPT) which will celebrate its 50th anniversary at the Spring 86 National ACS Meeting in New York. That committee of twelve sets standards for bachelor's degree programs in chemistry. Currently 580 chemistry departments are on the ACS/CPT approved list—200 with the Ph.D., the highest degree offered; 130, the MS; and 250, the BS/BA. Those 580 (of the 1,100 or so in existence) produce 10,000 (of the 11,000) bachelor's degreed persons every year. Also one should note that as those 580 have about 10,000, faculty the average rate is one bachelors in chemistry per chemistry professor per year (in CPT-approved chemistry departments).

The ACS/CPT revised its guidelines for approval of bachelor's degree programs in chemistry in October, 1983. In preparation for that revision a survey was made of a 112 department sample regarding computer support, computer sophistication of the (some 1,500) faculty, and computer-based enhancements to the practice of chemistry in their curricula. In addition CPT solicited 'appendices' to those guidelines from the technical ACS Divisions. Three of the nine such (published in May, 1984) were from Divisions whose members were already deeply involved with computers, namely: Analytical Chemistry, Chemical Information and Computers in Chemistry. Those appendices listed concepts their members felt should be treated somewhere in the four year program together with sample questions by way of calibration with respect to level and scope.

The ACS/CPT appendices reflect that CPT's members include researching chemists from major universities and industry as well as chemists from colleges. Thus CPT constitutes a forum overlapping research in chemistry and education of chemists. The appendices were generated by the corresponding technical divisions of the ACS: appended as an example is that produced by the ACS Division of Computers in Chemistry. The reaction to that appendix has been that much more is being called out for inclusion in the bachelor's degree program than might reasonably be incorporated—yet—spokepersons for the ACS Division of Chemical Education have recognized that appendix as a target to be worked and a goal to be achieved.

Recently there was formed within the ACS the Computer Secretariat whose membership includes six ACS Divisions as follows:

- Analytical Chemistry
- Chemical Education
- Chemical Information

- Computers in Chemistry
- Industrial and Engineering Chemistry (in Process)
- Physical Chemistry

Thus a formal computer-oriented coupling mechanism is coming into play within the ACS. In addition to better coordination of computer-related activities within the ACS, the *Computer Secretariat makes it feasible for the chemistry establishment to speak to the other* computer-focussed professions and policy makers with a united voice. Examples are implementation of the ACS ComSci Task Force on Large-scale Scientific Computing recommendations, and the application by the ACS to the American Federation of Information Processing Societies (AFIPS) American Federation of Information Processing Societies (AFIPS) for affiliateship (hopefully charter member status).

Computer-based electronic conferencing can reduce the intercommunication barrier between all the components of ACS as well as between ACS and other audiences. In particular, ACS has started a trial national network involving a few ACS Committees such as the Committee on Professional Training. In addition, with N.S.F. support, Project SERAPHIM based at Eastern Michigan University has instituted CHYMNET(tm) whereby participants in joint high school-college chemistry teacher computers-in-education summer projects can continue to communicate with each other during the academic year from their respective institutions geographically distributed all over the USA. Finally the concept of regional nodes dedicated to electronic conferencing, CHEMNET(tm), also has gained attention where the focus on local persons communicating with local persons facilitates easy communication between practicing scientists and science educators as well as making possible enhancements to local ACS Section activities.

The Division of Computers in Chemistry, created eleven years ago by the ACS, has conducted a large number of symposia—most of them published—dealing with some focussed area of computers in chemistry. A more recent activity of the ACS Division of Computers in Chemistry is a further illustration of an initiative from a technical division toward opening new vistas in chemistry education. At the April 85 ACS National Meeting in Miami Beach, Florida, the Symposium on Academic Programs in Computational Chemistry drew several papers illustrating a variety of responses to the overlap between 'computer science' and chemistry. Perhaps the most broadly based presentation was that by Professor MICHELE SCULLANEf Southeastern Massachusetts University (SMU). She described how they had augmented chemistry courses via the laboratory (in analytical, instrumental and physical chemistry) with both data reduction as well as on-line computer applications. Also they developed two chemistry courses in computer programming and numerical methods—one in the freshman year and one associated with physical chemistry.

Following that SMU instituted a minor in computer and information sciences (CIS) where the chemistry major, starting with the third semester, can take a CIS course every semester. The students in that CIS minor take about 130 semester hours and are still certifiable as having completed an ACS/CPT approved program whereas their colleagues who take the regular certifiable chemistry program need take only about 120 semester hours.

Graduates of the CIS-minor-augmented bachelor's degree in Chemistry program can go to graduate school in chemistry or in computer science.

The goals of that program are as follows:

1. Provide students with more flexible career options
2. Recruit more and better chemistry majors
3. Encourage chemistry faculty colleagues to broaden skills

4. Increase number of CIS minors

5. Make feasible *bailout* from the CIS minor at any time.

The most inspired part of SMU program is the creation of a new freshman course, a one-semester-hour seminar on computer science as a discipline. Students come to college full of excitement about entering the Information Age yet they have only vague and uneven impressions of just what is really going on. The canonical course in computer programming many colleges require of their incoming freshmen usually serves only to acculturate the student to the local computing environment—and even then the students do not make serious use of the computer until their upper division chemistry courses. With that seminar the students are able to see that they can major in an established substantive discipline and still get deeply into operating system, utilities and applications computer software as well as computer architecture.

The Division of Chemical Education, on the other hand, has an active Committee on Computers in Chemical Education (with a quarterly C.C.C.E. Newsletter now in its eighth year) as well as significant coverage of computers in education via articles in the ACS *Journal of Chemical Education*. (The ACS Journals Department, rather than the individual technical divisions, publishes the some 20 ACS technical journals—all with refereed articles. The *Journal of Analytical Chemistry* occasionally includes tutorials on aspects of computers in chemistry ranging from laboratory networking to scientific wordprocessors.) The C.C.C.E. conducts many regional and national work shops for high school and college chemistry teachers on computers in chemical education.

At the ACS National Meeting in Chicago in September 1985 the Division of Chemical Education is conducting a Symposium on Writing Across the Curriculum. A special effort is made via the Division of Computers in Chemistry to demonstrate the machine-based aids to analysis of English text coming to be available. Out of AT&T-Bell Laboratories has come the Writer's Workbench(tm), a collection of programs which analyse English text from the perspective of WILLIAM STRUNK and ELWYN BROOKS WHITE's *Elements of Style* [SRW-791]. Sentence parsing algorithms make further enhancements possible. Derivatives of that system now are available which run on common personal computers. The timing of those developments is quite good for chemistry curricula as the complaint most often heard by ACS/CPT about new bachelor's degreed persons in chemistry is that they have poor communication skills. Access to English text analysis tools means that chemistry professors can assign writing projects such that the mechanics of using the English language can be analysed by the machine and the chemistry professor can deal with the substantive part of the composition.

Of the 580 schools in the U.S. with ACS/CPT approved chemistry departments 125 have chemical engineering departments which must be served. That follows because the AIChE requires that undergraduate chemical engineering majors take their chemistry courses, through physical chemistry, from the chemistry department. In order to have good communication between those two groups there are occasional joint meetings; the most recent such happened in March of 1984 focussed on the content of undergraduate physical chemistry courses. That three-day retreat involved six physical chemists from industry and academe and six chemical engineers also from industry and academe. In addition PETER ATKINS and GILBERT CASTELLAN were present as they are authors of widely used physical chemistry textbooks. Called out for special attention at that Invitational Workshop was the impact of the computer. That workshop represents yet another cooperative effort between the several sectors of the chemistry establishment with a vested interest, namely; CPT and the Divisions of Chemical Education, Industrial and Engineering Chemistry and Physical Chemistry.

Another phenomenon affecting education of chemists because of the advent of the computer is that numerical methods now can be applied to solution of chemical problems at a much earlier level than hithero. The Mathematical Association of America has a panel on service courses which has been meeting with representatives of those disciplines whose majors take several lower division mathematics courses to determine just how the sequence and emphases of concepts should be revised. Last January two representatives of the ACS (a Chemometrician and a Theoretical Chemist) met with that panel. The (primarily numerical) methods identified for that MAA panel as being basic to the practice of chemistry were:

1. Numerical solution of transcendental equations

2. Simultaneous linear equations

3. Regression analysis

4. Numerical integration

5. Numerical solution of differential equations

6. Eigenvalue/eigenvector problems.

Further and more structured interactions are being planned.

There are some 7,000,000 known chemical substances whose syntheses have been described in the literature and that number is increasing at the rate of about 7,000 per week. Accordingly, undergraduate chemistry majors need to learn about corresponding computer-based chemical information resources. They need to become sensitive to the need to transform a three-dimensional simply-connected molecular diagram to a string of characters and to do that in a manner which facilitates searching through a corresponding data base. In order to address that issue, the ACS has had the Chemical Abstracts Service develop a College Plan whereby, on a sliding scale depending on size of the school and the kind of utilization, a substantial discount is available for both the printed version of Chemical Abstracts and CAS ONLINE (the latter outside of peak hours at a 90% discount). At this point in time there are over 240 active accounts from 173 participating schools. Also the ACS Division of Chemical Information—yet another ACS technical division—has prepared a comprehensive and current bibliography on literature searching for interested schools.

Up until recently the ACS had been the only major scientific discipline-focussed society in the U.S.A. with bachelor's degree approval program. In 1980 the ACS/CPT approval program was presented to the President of the Association for Computing Machinery as a model for a similar program for undergraduate programs in computer science. That act started a series of events which led to the creation of the Computing Sciences Accreditation Board in July, 1984. Thus standards are being introduced into undergraduate computer science education. That affects undergraduate chemistry education as many chemistry departments require their undergraduate majors to take a computer science course. Indeed, as noted above, a few departments of chemistry already have created a computational chemistry option or co-major.

A related event, following an initiative by a member of the ACS/CPT in 1981, was the creation by The College Board of an Advanced Placement Exam and corresponding course outline in Computer Science. That one-year college-level course was first introduced in fall, 1983esponding course outline in Computer Science. That one-year college-level course was first introduced in fall, 1983, and on 10.05.1984, 4,300 high school students from 915 high schools (of the 23,000 in the US) took the AP/CS Exam. Half again that number took the second such exam on 09.05.85 across the U.S.A. The number of participating high schools also increased (by 40%). PASCAL is the AP/CS language of choice. Thus standards are being introduced into computer science education in high schools and that is bound to affect the high school education in computer science of college-bound students (of the caliber desirable as 'feedstock' for the chemistry establishment).

Unfortunately, introduction of computer-based enhancements to chemical problem solving in the undergraduate program tend to happen primarily at the *upper* division level. Thus incoming freshmen, stimulated by their focussed computer science education and a general awareness they will live as part of the Information Age, are not likely to discover *in time to become chemistry majors* that chemistry is probably the richest scientific discipline with respect to the range and depth of computer use.

The most critical need with regard to attracting better students to chemistry careers at this point in time is articulation of the feeder schools with college chemistry curricula, particularly through incorporation of non-trivial computer-based enhancements to the practice of chemistry into the lower division-courses—especially the first year chemistry courses.

Chemists most prominent in research are the persons looked to as role models and for leadership in the practice of chemistry. They are the most visible exemplars of the intellectual achievements possible in that central—and expanding—science of chemistry. The education of future chemists is too important to leave to chemistry educators alone. Hence the **E** standing next to the **R** in **ICCCRE**.

# Computational Methods

# 2

# Computational Chemistry in Industrial Research

WOLFGANG T. DONNER
ZF-DID
ZF EDV-Entwicklung und Systeme
BAYER A.G.
D-5090 Leverkusen (DE)

## 2.1 Overview

The use of the computer for R&D problems is just one of many other computer applications in industrial companies. There are some differences between scientific and other applications like administrative data processing or process control. Within scientific applications there is a singularity in chemistry as the chemical structural formula contains many kinds ofinformation about molecules. Therefore the treatment of non-numerical information is in chemistry at least as important as classical numerical calculations.

The requirements of the synthetic chemist on an information system are strongly connected with his activities:

- He needs information about compounds—compounds that are already known, compounds that are available, compounds with common substructure, *etc.* Systems doing this are shortly mentioned. They differ from traditional numerical data retrieval systems as they are based on the treatment of graphical information rather than on data. As these systems have to perform some kind of chemical reasoning as soon as one asks for structurally related compounds, they tend towards intelligent assistents.

- He needs information to synthezise new molecules. Here reaction retrieval and computer assisted synthesis planning are the wanted tools. Among synthesis planning programs we find early examples of expert systems.

- He needs molecular information as electron densities, reactivities, conformations *etc.* These data can be obtained by quantum chemical or force field calculations, or experimentally from spectroscopy or structure elucidation.

- He obtains data for his substances when they are tested for biological activities or other properties. His endeavour to come to better activities or properties by modifying the molecular structure can be supported by correlation methods and the use of empirical or calculated parameters. The aim is to find a link between chemical structure of a compound and properties of the substance.

- Graphical manipulation of chemical structures in three-dimensional space is another tool that offers a new approach to the understanding why certain molecules act in a specific way, while others are inactive.
- He has to write reports on the results of his research work and finally , bringing together molecular structures, test data, and text. For that purpose word processing systems are needed.

For the chemist who in most cases is not an expert in data processing, these components should be parts of a uniform and integrated system, *i.e.* a system with a uniform user interface, which can be used easily even by the occasional user. The structures the chemist once has stored within the system or selected from a data base of structures should be available for any program within this system.

Other systems exist or will come to existence for structure elucidation (analytical chemistry), for clinical or biological tests, for measurements of properties of substances, *etc.* They all will have to communicate with each other, either by using the same data bases or—more often—by having at least common data structures.

Aspects of future developments and their possible consequences for the bench chemist are finally sketched.

## 2.2 Introduction

As in many other industries the computer is widely used today in chemical industry. One will hardly find any division within a company where the computer has not found its place as an important tool. This goes from rationalization of personnel intensive labour to the creation of new qualities by using the specific merits of the computer.

Often the applications of the computer are categorized in the following way:

- Administrative Data Processing
- Process Control
- Scientific Applications

This paper is concerned with the third point—scientific applications. However, there are many problems in this field that are common to those in other applications. To mention a few:

- The antagonism between centralized and decentralized data processing.
- The antagonism between general purpose computing and dedicated or even personal computing.
- The antagonism between different data base management systems (hierarchical *vs.* relational DBMS).

From other aspects we find some criteria that help to differentiate between the needs in these three fields:

In many cases data processing is used for structured problems, this holds for administrative data processing as well as for process control. In such cases it is rather easy to compare costs and benefits. Problems in research are often not structured. Here, other aspects than cost-benefit considerations have to be considered to find arguments for using the computer. Even in cases like the online use of data bases, one cannot compare and rate the gains (completeness of information, fastness of obtaining the information) against the costs. Despite