(106) Symposium on Computer-Assisted Organic Synthesis, Division of Chemical Information and Division of Computers in Chemistry (joint sponsors), 171st National Meeting of the American Chemical Society, New York, N.Y., April 4–9, 1976.
(107) R. Fugmann, W. Braun, and W. Vaupel, "GREMAS-ein Weg zur Klassification und Dokumentation in der organische Chemie", *Nach. Dok.*, **14** (4), 179–190 (1963).
(108) R. Fugmann, H. Nickelsen, I. Nickelsen, and J. H. Winter, "A Topological Method for the Representation of Synthetic and Analytical Relations of Concepts", *Angew. Chem., Int. Ed. Engl.*, **9**, 589–595 (1970).
(109) R. Fugmann, H. Nickelsen, I. Nickelsen, and J. H. Winter, "Representation of Concept Relations Using the TOSAR System of the IDC: Treatise III on Information Retrieval Theory", *J. Am. Soc. Inf. Sci.*, **25** (5), 287–306 (1974).
(110) M. Osinga and A. A. Verrijn Stuart, "Documentation of Chemical Reactions. I. A Faceted Classification", *J. Chem. Doc.*, **13**, 36–39 (1973).
(111) M. Osinga and A. A. Verrijn Stuart, "Documentation of Chemical Reactions. II. Analysis of the Wiswesser Line Notation", *J. Chem. Doc.*, **14**, 194–198 (1974).
(112) J. E. Armitage and M. F. Lynch, "Automatic Detection of Structural Similarities among Chemical Compounds", *J. Chem. Soc. C*, 521–528 (1967).
(113) J. E. Armitage, J. E. Crowe, P. N. Evans, M. F. Lynch, and J. A. McGuirk, "Documentation of Chemical Reactions by Computer. Analysis of Natural Changes", *J. Chem. Doc.*, **7**, 209–215 (1967).
(114) R. Clingling and M. F. Lynch, "Production of Printed Indexes of Chemical Reactions. I. Analysis of Functional Group Interconversions", *J. Chem. Doc.*, **13**, 98–102 (1973).
(115) R. Clingling and M. F. Lynch, "Production of Printed Indexes of Chemical Reactions. II. Analysis of Reactions Involving Ring Formation, Cleavage, and Interconversion", *J. Chem. Doc.*, **14**, 69–71 (1974).
(116) J. Valls, "Reaction Documentation", in "Computer Representation and Manipulation of Chemical Information", W. T. Wipke, S. R. Heller, R. J. Feldmann, and E. Hyde, Ed., Wiley, New York, N.Y., 1974, pp 83–103.
(117) J. R. Koskinen and B. R. Kowalski, "Interactive Pattern Recognition in the Chemical Laboratory", *J. Chem. Info. Comput. Sci.*, **15**, 119–123 (1975).
(118) P. D. Jurs, "Chemical Data Interpretation Using Pattern Recognition Techniques", "Computer Representation and Manipulation of Chemical Information", W. T. Wipke, S. R. Heller, R. J. Feldmann, and E. Hyde, Ed., Wiley, New York, N.Y., 1974, pp 265–285.
(119) C. Hansch, A. Leo, and D. Elkins, "Computerized Management of Structure-Activity Data. I. Multivariate Analysis of Biological Data", *J. Chem. Doc.*, **14**, 57–61 (1974).
(120) S. R. Heller, R. J. Feldmann, H. M. Fales, and G. W. A. Milne, "A Conversational Mass Spectral Search System. IV. The Evolution of a System for the Retrieval of Mass Spectral Information", *J. Chem. Doc.*, **13**, 130–133 (1973).
(121) D. J. Floto, R. A. Hagstrom, T. L. Heying, R. E. Maizell, and W. J. Mayer, "Technological Forecasting—An Experiment Relating to the Pulp and Paper Industry", *J. Chem. Doc.*, **13**, 162–165 (1973).
(122) D. H. Smith, L. M. Masintes, and N. S. Sridharan, "Heuristic DENDRAL: Analysis of Molecular Structure", "Computer Representation and Manipulation of Chemical Information", W. T. Wipke, S. R. Heller, R. J. Feldmann and E. Hyde, Ed., Wiley, New York, N.Y., 1974, pp 287–315.
(123) M. F. Lynch, J. M. Harrison, W. G. Town, and J. E. Ash, Ed., "Computer Handling of Chemical Structure Information", Macdonald, London, 1971.
(124) W. T. Wipke, S. R. Heller, R. J. Feldmann, and E. Hyde, Ed., "Computer Representation and Manipulation of Chemical Information", Wiley, New York, N.Y., 1974.
(125) J. E. Ash and E. Hyde, "Chemical Information Systems", Ellis Harwood, Ltd., Chichester, England, 1975.

Status of Computers and the Information Industry and Potential Future Trends†

CARLOS M. BOWMAN* and PATRICIA F. ROUSH

Central Research–Computation Laboratory, The Dow Chemical Company, Midland, Michigan 48640

Received April 30, 1976

Computer technology has advanced at a tremendous rate in the last three decades. Its influence has had a very significant impact on chemical information science. Large amounts of information are now available at the fingertips of the user through interactive systems. Data acquisition systems now permit the direct recording of experimental data from the laboratory. The days of the printed journal are numbered, and the use of computers in the planning of experiments will soon be a reality.

It is truly a pleasure and an honor to participate in this symposium with such distinguished speakers, and particularly in the first Herman Skolnik Award Symposium. It has been my (C.M.B.) privilege to know Herman for many years. I have come to respect his grasp of the important aspects of information science and his perceptive evaluation of work in this most important field. The naming of an award for him is truly fitting. The object of this talk is to tell you about computers in chemical information.

It is not surprising that computers and computer technology have had a tremendous impact on chemical information science as they have on almost every other aspect of our life. It is our intention to sketch briefly the development of computer technology over the last three decades and to try to relate this development to what we consider the golden years of chemical information science. We will then attempt to suggest some areas where new developments will take place and what some of the problems facing us will be.

Computer technology has developed in three major areas: processing capability, storage capacity, and real time interactive systems.

Processing capability has grown from operational speeds of a few seconds to a few nanoseconds. That is an improvement of almost nine orders of magnitude. Storage capacity has increased from being almost nonexistent, when data were stored on cards, to a very limited amount of magnetic drum and core storage, to solid-state storage which now permits immediate access to several million characters. One of the key developments in storage capacity was the introduction of magnetic media, including tapes, followed by disks. There is virtually no limit to the amount of information that can be put on-line in a computer system today.

Most computer usage in the early days was on the batch basis. It has only been seven or eight years since interactive systems began to make significant in-roads in this area. Today, a large portion of all computing is done in an interactive or

† Presented at the Award Symposium on "Contributions of the Division of Chemical Literature (Information) to the Chemical Society" at the ACS Centennial Meeting, New York, N.Y., April 6, 1976.

time-sharing mode. In addition, the ability to interface directly with instruments, and the ability to capture data as it is being produced, eliminates many transcription and input problems.

The increasing ability to process more and more material, at greater speeds, to add to storage capacity, and to be able to access information in a time-sharing environment has made computing an integral part of almost every phase of a chemist's life. One very important development which will not be discussed at this point is the miniaturization of computers and the development of minicomputers and microcomputers.

What effect has the tremendous increase in processing capability had on the handling of chemical information? Some of the very early applications of computers to chemical information were in the production of tabulated data, usually physical properties, bibliographies, and other such tabulations. The computer was nothing more than a typewriter or an easy way of arranging printed matter in various forms. As computing capacity increased, it became apparent that the information could be manipulated as it was being tabulated or printed. One of the earliest forms of information processing was the Key-Word-In-Context index. This was a first attempt at automatic indexing. Today, this tremendous computing capacity allows us to search through large amounts of text to extract pertinent sentences or paragraphs, and to answer questions directly rather than merely pointing to a document or a reference.

As storage capacity has grown, the amount of information that can be stored and retrieved by computer has been greatly increased. Some of the very early files were kept in drawers of cards. In fact, one of the first projects in this field at Dow was to develop a computer system that would eliminate the need to manipulate hundreds of thousands of punched cards. Magnetic tapes permitted the storage of bibliographic information and various types of indexes. However, the advent of direct access devices such as disks and drums permitted the development of systems using techniques such as Uniterm indexing with inverted files. This opened a way for direct access systems. The ability to store literally billions of characters on-line has resulted in the development of the data base concept, which allows for a large body of similar information to be stored in an organized manner which can then be accessed by more than one person at a time and in a variety of ways.

The most frustrating and difficult part of applying computers to information science has been to try to bridge the gap between the scientist and the machine. Although we have been very successful at developing complex and efficient information systems, and although we can manipulate billions of characters in a fraction of a second, we are as yet unable to communicate and interact easily with a computerized information system. Most systems that have been devised to do this are still very cumbersome and require a great deal of structuring of the input and in-depth knowledge of the system in order to be used efficiently. Recent developments in this area, which will be discussed later, show some hope.

So we see that computers have come a long way in the last 25 years and have become deeply enmeshed in the operation of information systems. Let us turn now to more specific areas where computers have had an impact, and examine them as information flows from its source to its use.

The generator and ultimate user of chemical information is the chemist. He creates information in the laboratory and, to plan his work, he goes to the literature to find out what has been done. Until recently, computers had very little direct impact on the chemist. A revolutionary development in the computer industry permitted the miniaturization of computing systems and the manufacture of small, very powerful and rather inexpensive computers. Minicomputers have literally transformed the scientist's work place. He can control his experiments by the use of the minicomputer. He can accumulate his data in this device and perform fairly complex operations to produce printouts, plots, or graphs, or he can transmit the data to much larger and more powerful systems for further analysis. The number of instruments that can be interfaced to minicomputers is almost limitless. There are many chemists in colleges and universities who are now learning, as an integral part of their training, the programming of these computers, and in many cases the actual building and design of interfaces to link instruments with computers. A number of commercially available instruments have built into them the computer facility to allow for capture of data and control of experiments. Thus we see that the computer is now involved in the generation of information. We have not yet advanced, of course, to the point where the computer will take the data and write up a paper. However, the data can be generated in a much more timely and efficient manner than ever before.

One other area where computers have had a tremendous influence on the chemist is in planning his research. There are a number of systems available to the organic chemist today that will allow him to explore on paper different routes for the synthesis of a new compound or material. This has required the development of one of the most interesting and fascinating applications of computer technology to chemical information, that is, the identification and manipulation of chemical structures. Chemistry is unique in that it has developed a very precise means by which to characterize the materials with which it works. Thus a structural diagram of a molecule is so precise that it lends itself to manipulation, storage, and comparison by computer systems. A number of systems have been developed to store such information, and a number of computer programs have been written that manipulate chemical structures. A significant amount of the effort in chemical information has revolved around the development of techniques for the identification, storage, and retrieval of structural diagrams.

Let us try to follow the flow of written information as it goes through the cycle and examine the impact that computers have had on this flow. The chemist, once he has written the rough draft of a paper, can now have it typed on a typewriter that has a small computer associated with it which will permit easier editing and proofreading. The manuscript then goes to an editor who could have a computerized system to keep track of the many manuscripts that he has and their progress through the selection process. Once the manuscript has been accepted for publication, it undoubtedly will be processed by an operation which will have a computer as an integral part of the driving mechanism for the presses. It could be computer typesetting, computer driven photocomposition, or many of the other techniques available today. The production of the journal article is in most cases now directly integrated with a computer system. This includes an on-line editing system.

Once the paper has been produced, the next step is the abstracting and indexing as carried out by Chemical Abstracts Service. It should suffice to say that every aspect of *Chemical Abstracts* involves the application of computers and computer technology. Computers are used to edit the abstract, to identify compounds, to produce the indexes, and to group the abstracts in specific groupings so that they can be published on a selective basis. Ultimately this information finds its way into a data base that can be searched on-line. *Chemical Abstracts* is an excellent example of the application of computers to the information industry.

Having produced the journal and the abstract, we turn now to the general repositories of information, the libraries and information centers. Computers are now being used in libraries to control the inventory and the flow of books and journals and to search the literature. For example, much of the information generated by *Chemical Abstracts* in this decade is now available at a nominal cost to any individual who has access to a computer terminal and a telephone. Libraries and information centers are the primary users of these data bases mainly because of the complexity of their use. The occasional user cannot address these types of systems and obtain satisfactory results, as he needs to be thoroughly familiar with the methodology of the system in order to obtain the desired information.

We then have come a full cycle, from the chemist in the laboratory that is generating information using his minicomputer, back to the library or center where the chemist accesses the information. In every step of this cycle we find computer systems as integral parts of the process.

One other area which is not usually thought about when we think of chemical information, but which is receiving a great deal of attention today, is the development of computer systems and programs which will allow us to interpret the meaning of the information and to make inferences from it. The development of so-called "artificial intelligence" techniques has greatly broadened the horizons of the information chemist.

Let us now turn our attention to some future trends and what our information systems may be like in the decades to come.

The significant increases in computing power and capacity over the last decades will not be continued. There will be a leveling off in the growth rates for computing power, storage capacity, and machine size. The emphasis is now and will continue to be on providing these capabilities at lower and lower cost, thus making them available to more people. The trend started by minicomputers will continue as we have seen more recently with the advent of microcomputers. This is all in the effort to provide computing at a very low cost.

Another significant development that is already in the offing but will certainly become a practical reality is extensive computer networks that will allow individuals to use many computer resources and share information with their colleagues throughout the world. Systems will become easier and easier to use, and the chemist of tomorrow will be more familiar with computers and will learn to use them to a much greater extent than many of us today.

It is not improbable that the dream of the science fiction writer which portrays a chemist in a laboratory with a screen and keyboard as his main tools will become a reality. He will be able to access all the information of the past and extract from data bases anywhere in the world the pertinent material that he needs. He will then apply the necessary manipulation and transformations to that data to test his hypothesis and theories, such that when he finally goes to the laboratory bench to actually perform an experiment, it will be so well thought out, and so well designed, that he will derive a maximum amount of information from it. The experiment will be carried out on a well-controlled basis with all the data being captured directly for him and returned to his computer system to compare with his original plans. This will allow the chemist to utilize his personal, God-given computer to the utmost.

The results of these well-planned experiments will then become a part of the data base, thus becoming available to all. This could well circumvent the great publishing and abstracting efforts. Timeliness and quantity will cease to be a problem, but issues of privacy and accessibility will come into prominence.

It sounds improbable, but it is the future and we must be prepared to deal with it.