indicate when and how this flag is set. The recovery of a tautomeric input structure is obtained by using these rules in reverse. (Where the original structure is no longer known, e.g., if the input was obtained from a system in which structures are normalized, this can be indicated by placing the number 3 into the TAU category).

The original structures of other types of compounds can be recovered in a similar manner. For example, cyclic dienes can be specified by retaining the location of one original double bond for each ring system. The odd ring atom (Figure 5) can be specified by a special flag, comparable to TAU. But because there was little need for these additional recovery capabilities, they are not available in the WRAIR system.

For retrieving compounds with shifting structure, a user of the WRAIR system thus has the option of searching either under relaxed or under stringent identity rules. There is thus less need for caution in the application of tautomer codes than there is with the use of normalized bonds. Extension of the above method to keto–enol tautomers may consequently be considered in the future.

5. ACKNOWLEDGMENT

I am indebted to Ted N. Fujimoto for programming support and to Ms. Margaret Milne, of the University of Pennsylvania, for her invaluable comments and suggestions.

REFERENCES AND NOTES

(1) L. Hodes and A. Feldman, "An Efficient Design for Chemical Structure Searching. II. A Solution to the Large Data Base File Structure Dilemma", in preparation.
(2) G. W. Adamson, M. F. Lynch, and W. G. Town, "Analysis of Structural Characteristics of Chemical Compounds in a Large Computer-Based File. II. Atom-Centred Fragments", *J. Chem. Soc. C*, 3702 (1971). D. J. Gluck, "A Chemical Structure Storage and Search System Developed at Du Pont", *J. Chem. Doc.*, **5**, 43 (1965); M. Milne, D. Lefkovitz, M. Plotkin, H. Hill, and R. Powers, "A Study of the Potential for Useful Exchange in Chemical Structure Handling Areas of Four Chemical Information Systems", and references cited therein, available from NTIS, AD 782, 239.
(3) H. L. Morgan, "The Generation of a Unique Machine Description for Chemical Structures-A Technique Developed at Chemical Abstracts Service", *J. Chem. Doc.*, **5**, 107 (1965).
(4) A. Feldman and L. Hodes, "An Efficient Design for Chemical Structure Searching. I. The Screens", *J. Chem. Inf. Comput. Sci.*, **15**, 147 (1975).

An Interactive Computer Graphics System for Processing Chemical Structure Diagrams†

JAMES E. BLAKE, NICK A. FARMER,* and REGINALD C. HAINES
Chemical Abstracts Service, P.O. Box 3012, Columbus, Ohio 43210

Received February 3, 1977

An interactive computer graphics system has been developed at Chemical Abstracts Service (CAS) which allows creation of chemical structure diagrams of a quality suitable for publication in *Chemical Abstracts* (CA). The system is based on a Digital Equipment Corporation (DEC) PDP-15 graphics system connected to an IBM 370/168 host computer. Employing a light pen, the user at a terminal creates a structure diagram by selecting individual atoms or chemical rings from a menu list and by indicating how the various pieces are to be connected. The computer program then calculates the preferred placement of the structural pieces according to a set of formatting rules. The user at the terminal can, however, override these rules by moving items on the screen with the light pen. Once input, structure diagrams are stored on disk (or tape) at the IBM 370 until required for insertion into a specific publication. Then text and graphic data are selected from the IBM 370 database and photocomposed on an Autologic APS-4, producing full-page publication quality output. This automated system has been in daily use for over two years. Two reference files have been built in this manner. One file includes basic chemical ring shapes and currently contains over 19 000 entries. The second file, which currently contains over 42 000 structure diagrams, is the source for the structure diagrams that routinely appear in the CA Volume and Collective Chemical Substance Indexes and in the Parent Compound Handbook. The system is also used to produce the structure diagrams that appear in the weekly issues of *Chemical Abstracts*.

INTRODUCTION

Each year Chemical Abstracts Service (CAS) publishes over 59 000 pages of chemical information, consisting mainly of the weekly issues of CA and its volume indexes, which are published every six months. (These page number figures do not include the Collective Indexes to CA which are published every five years.) These publications represent over 800 000 000 characters of textual information and more than 62 000 chemical structure diagrams. Although the first steps toward the computer-controlled composition of CAS publications began in 1965, it was not until 1971 that attention was directed toward the computer processing of graphical data. The computer-composed publications had "windows" in the text where the graphics were to appear. Graphics were drawn by illustrators, photoreduced, and then stripped (pasted) into the page containing the text that had been photocomposed. Over the past five years CAS has developed a system that provides the capability for computer composition of graphical as well as textual data. This paper describes a part of that system. The first section describes the input sub-system, its hardware and software structure, and some of the basic system characteristics. The next section describes three applications which process chemical structure diagrams that have been implemented using this system. The last section summarizes our experience gained from using this system and describes our plans for future work in this area.

There have been several computer graphics systems developed over the last few years that process chemical structure diagrams. The system developed by Corey and Wipke is used for computer-assisted synthesis of organic compounds.[1] The PROPHET system, developed by Bolt, Beranek, and Newman, Inc., is used for information retrieval in a medical environ-

† Presented in part at the 169th National Meeting of the American Chemical Society, Philadelphia, Pa., April 8, 1975.

* Author to whom correspondence should be addressed.

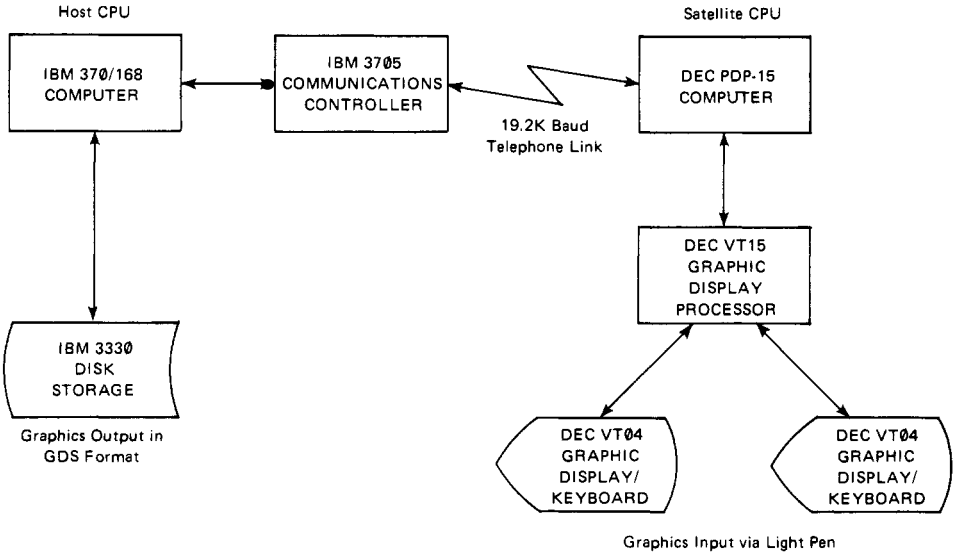


Figure 1. Graphics system hardware.

ment.[2] Several substructure search systems, such as those developed by Feldmann[3] and more recently by Kowalski,[4] use computer graphics to process chemical structure diagrams. The system described in this paper is used for a much different purpose, namely, providing "graphic arts quality" structure diagrams suitable for use in a printed publication.

SYSTEM DESCRIPTION

The hardware on which the CAS Interactive Computer Graphics System is implemented consists of two central processors (CPUs) and a variety of input/output (I/O) devices. A block diagram of the system is shown in Figure 1. The input and display of graphical data are handled by the satellite CPU, a DEC PDP-15 minicomputer with 24K 18-bit words of memory.[5] Two graphics terminals, each consisting of a refresh-type cathode-ray tube (CRT), an alphanumeric keyboard, a light pen, and six push-buttons, are connected via a DEC VT15 display processor to the PDP-15. A 19.2-K bit telephone line connects the satellite with the host CPU, an IBM 370/168 with 3 million bytes of memory. The host CPU handles the on-line storage and retrieval of graphical data on disk and is also responsible for most of the computations and manipulations performed on these data.

Two distinct pieces of software (both written in basic assembler language) have been developed at CAS for handling on-line graphical data. The first system developed was the CAS Interactive Computer Graphics System (GS), which is a general-purpose interactive computer graphics system containing the basic functions required for manipulating any kind of graphical data. The second system developed was an application system called the On-Line Structure Input System (OLSIS). The basic function of OLSIS is to interact with the user at the terminal, providing assistance in the input of chemical structure diagrams and structure-related information. Before either of these systems was developed, a device-independent Graphical Data Structure (GDS) had to be defined. This section of the paper will first define the format of the GDS and then discuss in detail some of the features of the GS and OLSIS.

THE GRAPHICAL DATA STRUCTURE

The design of the CAS Graphical Data Structure was based on three objectives:

1. The data structure must be readily manipulated in a real-time environment in order to provide the on-line user with realistic response time.
2. The data structure must be as compact as possible so as to minimize the amount of on-line disk storage required.
3. The data structure must be device independent inasmuch as the lifetime of the GDS files is certain to exceed that of any specific hardware device.

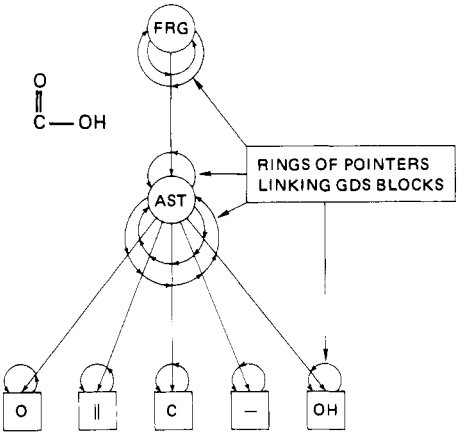


Figure 2. Structure diagram and GDS tree.

The data structure developed at CAS for manipulating and storing graphical data is in the form of an oriented tree.[6] Each tree is built out of four basic building blocks:

1. **NODE blocks**—provide the hierarchy in the picture. A specific node is associated with each subpicture. The OLSIS application system has 13 different node types. A node is represented graphically by a circle: (N1).
2. **BRANCH blocks**—connect nodes to nodes or nodes to leaves. Each branch contains the coordinates, intensity, scale, and other information affecting the subtree below it. A branch is represented graphically by an arrow: B1.
3. **LEAF blocks**—contain all of the characters and lines which describe the picture. The OLSIS application system has 18 different leaf types. A leaf is represented graphically by a rectangle: [L1].
4. **DATA blocks**—contain the nondisplayable, application-dependent information. Each of the previous three blocks may have an associated data block. In OLSIS the data blocks contain all of the chemical information that describes the chemical structure diagram, such as atom and bond types and atom by atom connectivity values. Data blocks are not shown in a graphical portrayal of the GDS.

Figure 2 shows an example of a chemical structure diagram and the corresponding GDS tree that would be generated by

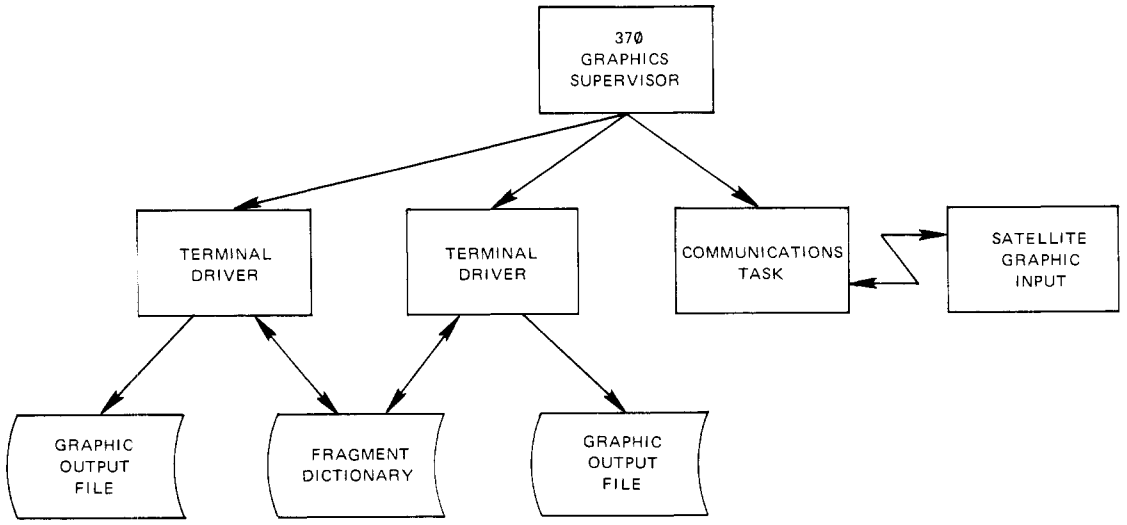


Figure 3. Graphics system multitasking architecture.

OLSIS. Note that since it is an oriented tree (by definition), the order in which the branches are attached to a given node is not significant. All of the blocks in the diagram are connected by circularly linked lists called rings. In addition, each atom and bond leaf has a data block (not shown) and these data blocks are also connected by rings. It is these latter rings that provide the information structure which indicates which atoms are bonded to each other in the chemical sense.

THE GRAPHICS SYSTEM

The CAS Interactive Computer Graphics System is a general-purpose, application-independent, interactive computer graphics system. Basically, the GS is composed of two sets of parallel programs, one written for the PDP-15 and one for the IBM 370. All the programs written for the 370 are device independent, except for a few modules that actually communicate with the PDP-15. The 370 GS was designed so that an application programmer can easily develop programs for the 370 without being concerned about the particular hardware peculiarities of the actual satellite CPU being used. The application programmer uses the GS by simply issuing subroutine calls to the appropriate GS command modules. There are two basic types of GS commands:

1. Noninteractive commands for creating and manipulating the GDS. Included among these are commands to allocate and initiate the basic blocks and commands to build, scan, and restructure GDS trees.
2. Interactive commands that activate and accept data from the keyboard, light pen, and push-buttons at the terminal. The light pen and push-buttons are used to perform such commands as (a) selecting special leaves or subtrees on the CRT, (b) moving leaves or subtrees to another portion of the screen, and (c) drawing lines on the CRT in a freehand mode.

The 370 GS has a multi-tasking structure; see Figure 3. The Graphics Supervisor Task controls the attaching and detaching of its Terminal Driver subtasks. There is one Terminal Driver Task for each terminal that is actually logged-on; this task contains the application-dependent program being run on that terminal. The Communications Task handles all of the GS commands being issued by the Terminal Drivers and performs all the I/O with the satellite graphics processor. It is possible to run the same reentrant application program on each terminal or to run a different one on each. It is important to note that in order to upgrade this system to a different satellite CPU, the only 370 modules requiring modification would be the I/O modules in the Communications Task. None of the extensive application code in the Terminal Drivers would be affected by such a hardware change.

THE ON-LINE STRUCTURE INPUT SYSTEM

The On-Line Structure Input System is the first application system developed at CAS to use the GS software just described. The basic purpose of OLSIS is to provide the user at the terminal with the capabilities needed to input chemical structure diagrams into the CAS computer system. Several requirements were placed on the design of OLSIS:

1. It should be able to handle any chemical substance processed by the CAS Chemical Registry System.[7,8]
2. It should produce chemical structure diagrams of graphic arts quality.
3. It should dynamically update all chemical information as the structure is being built, editing for chemical errors where possible.
4. It should be as interactive as possible, performing most of the routine work, with the user specifying only the command sequence required to build up the structure.

OLSIS has a well-defined hierarchical data structure for the chemical structure diagrams that it creates. This data structure contains all the necessary chemical information to correctly identify a completed structure diagram. A formal grammar can be defined to describe chemical structure diagrams; however, such a description would be too lengthy for the purposes of this paper.[9] Figure 4 illustrates the complexity and the hierarchy of the OLSIS data structure. The node types depicted in this figure are defined as follows:

STR	Total Structure
XPR	Expression Level
FRG	Fragment Level
SST	Substructure Level
CST	Cyclic Structure
AST	Acyclic Structure
ATM	Atom Level

Method of Input. In order to illustrate how OLSIS may be used to input a chemical structure diagram, it is first necessary to describe the OLSIS "menus" and their uses. The retrieval and use of substructures from the Fragment Dictionary will then be explained. Finally, the use of some of the structure building functions will be described.

Structures are input by selecting a sequence of commands from various menus; the command sequence chosen will

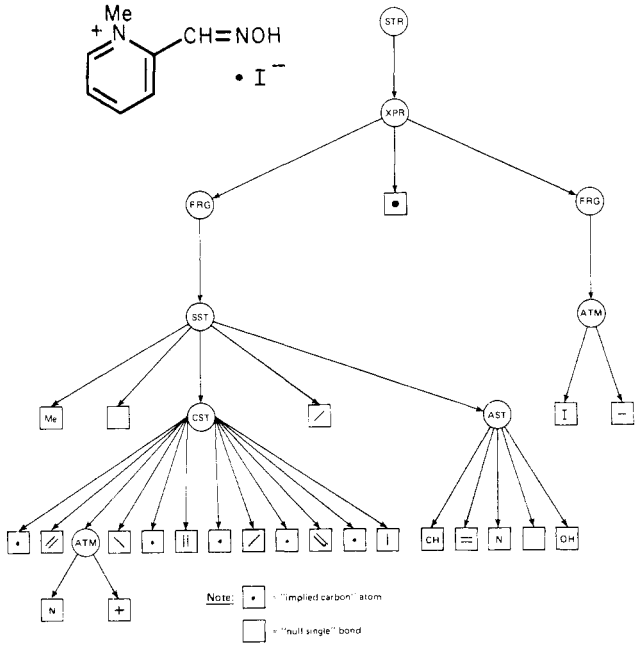


Figure 4. Hierarchy of OLSIS data structure.

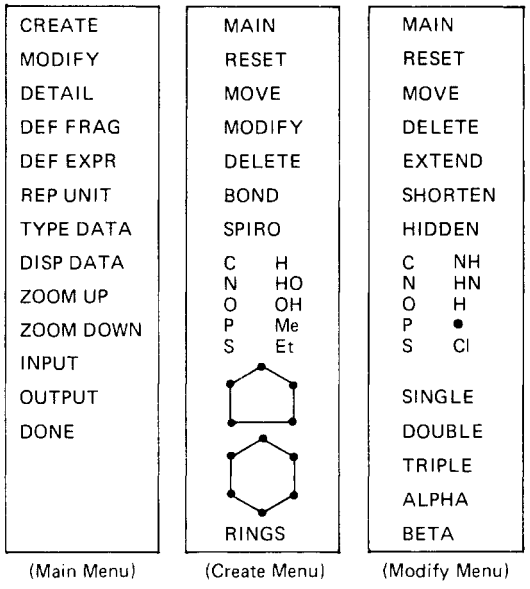


Figure 5. Examples of OLSIS menus.

determine the various components of the structure and how they are connected. These menus are simply lists of commands and structure components (such as atoms and rings) arranged in logical groups; they will appear (one at a time) in the right-hand portion of the screen. Individual items or commands on the menus are selected by using the light pen while building up a structure on the left-hand portion of the screen. The most frequently used menus are illustrated in Figure 5. The various menu items can be divided into three categories:

1. Items which pass control to another menu. The MAIN menu is always the first menu to appear after log-on. All other menus contain an item which will transfer control back to MAIN; thus one can always travel from one menu to another via MAIN.
2. Items which specify commands or functions to be performed on the structure being built. Items such as BOND, MOVE, and DELETE are examples of such commands.
3. Items which are chemical components or substructures. Element symbols, shortcut symbols, bond types, and rings are in this category.

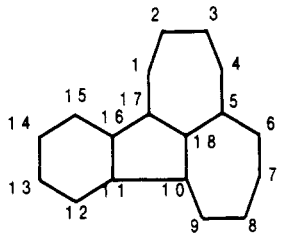


Figure 6. Typical ring graph diagram.

Since many structural fragments appear frequently in chemical structures, an on-line disk file called the Fragment Dictionary has been built which contains several thousand frequently occurring structural fragments. These fragments may be retrieved from the Fragment Dictionary via keyboard commands. After the name of the desired fragment is keyed, OLSIS will retrieve it and display it in one of two areas on the CRT. The right-hand portion of the screen, adjacent to the menu area, is called the "holding area" and may contain several temporary work structures. The left portion of the screen where structures are built up is termed the "main structure area." If the main structure area is empty, a fragment may be copied into it from the holding area by simply pointing the light pen at any atom in the fragment.

The usual technique for building new structure diagrams can be described as an interactive assembly of primitive substructure components. The process usually begins by the retrieval of one or more fragments as indicated above. These fragments are then combined and augmented by using the commands on the CREATE menu. For example, a bond can be created between two atoms by first selecting each atom with the light pen and then selecting the BOND command. OLSIS calculates the relative position of each substructure as it is assembled by the user. The result is a complete structure skeleton composed of standard ring shapes and uniformly placed heteroatoms.

After the basic graph of the structure has been constructed, the user then displays the MODIFY menu. The commands on this menu are used to change the atom and bond values as desired. For example, one or more atoms may be changed by first selecting a new element value from the menu and then selecting those atoms to be modified. In a similar manner, bond types may be changed to double, triple, etc. When the user has completed the chemical structure diagram, it is then written on a disk output file and the screen is cleared in preparation for the next structure.

Uses of the OLSIS System. At present, three distinct uses of the OLSIS system have been implemented at CAS. The first use of OLSIS has been in building a computer-readable file (called the Ring Image File) which contains the graphical representation for all of the ring graphs identified by the CAS Chemical Registry System. [The ring graph is the basic geometrical skeleton of a ring system. The types of atoms (e.g., carbon, oxygen, nitrogen) and the types of bonds (i.e., single, double, triple) are not defined.] The primary use of this file is to provide the basic building blocks used in the algorithmic generation of structure diagrams from the Registry Connection Tables.[10] The Ring Image File currently contains over 19 000 different rings and grows at a rate of about 30 new graphs each week. In the building of the Ring Image File, a hardcopy drawing of each graph was prepared by a chemist showing the proper orientation of the graphs. The graphs were then transcribed into computer-readable form using the OLSIS system. The connection tables associated with the ring graphs were then verified against the corresponding connection tables in the Registry System to ensure that no errors were made in the drawing or transcription processes. Figure 6 shows a typical ring graph.

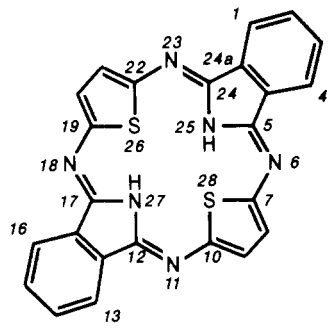


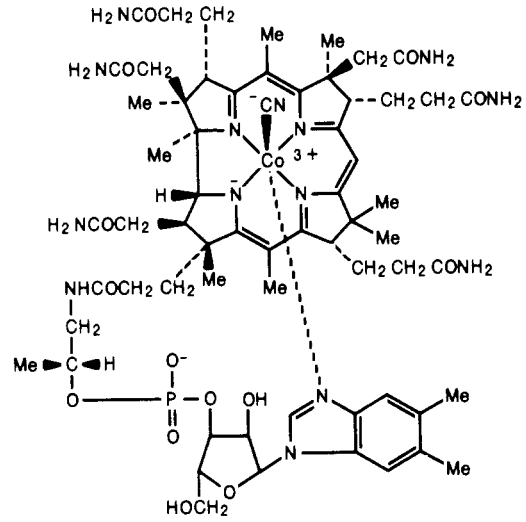
Figure 7. Typical structure cut diagram.

The second use of OLSIS was in building another computer-readable file (called the Structure Cut File) which contains the graphical representation for all of the structure diagrams that appear in the CA Chemical Substance Index. The Structure Cut File currently contains over 42000 structure diagrams, with about 60 new entries being added each week as the result of the identification of new information in literature covered in *Chemical Abstracts*. There are four types of structures in this file: metallocenes and boron cages; inorganic acids and their derivatives; stereoparents, which are chemical entities with common names that imply stereochemistry and are used as CA index heading parents; and ring parents, which are rings with particular bond variants chosen to represent families of ring systems showing the same ring skeleton and heteroatom positioning. The manual file of structure cuts was used as the starting point for building the computer-readable version. Each cut was first reviewed by a chemist and then transcribed into computer-readable form using the OLSIS system. The connection tables associated with the structure cuts were then verified against the corresponding connection tables in the Registry System to ensure that no errors were made during the transcription process. Figure 7 shows a typical structure cut. Computer-composed structure diagrams first appeared in the Volume 81 CA Chemical Substance Index. Since the building of the computer-readable Structure Cut File was not complete at that time, only about 60% of the structure diagrams were computer-composed, with the remaining 40% being processed by traditional manual methods. About 80% of the structure diagrams in Volume 82 were composed, and by Volume 83 this figure rose to 99.9% as the file building was complete. [The only structure diagrams not computer-composed are those that are too large to fit in the memory of the PDP-15, or those that have curved lines.]

The third use of OLSIS was in processing the structure diagrams that appear in the weekly issues of CA. On the average, about 1000 structure diagrams appear in CA each week. About 200 of these are commonly occurring structures which are automatically extracted from a computer-readable file of about 600 structure diagrams built expressly for this purpose (see Figure 8). The remaining 800 structure diagrams are input using OLSIS. Computer-composed structures began appearing in CA in February 1976. By April 1976, over 98% of all of the structure diagrams in CA were computer composed.

CONCLUSION

Computer composition of graphical data (in this case, chemical structure diagrams) requires a great deal of sophisticated computer hardware and software. However, our experience has been that, overall, computer composition of chemical structure diagrams can be done less expensively than composition done with traditional manual methods. Figure 9 shows some of the rates that we have experienced with the three applications mentioned previously. Note that these data



I, u414 Vitamin B_{12}

Figure 8. Repetitive CA structure diagram.

Use	Manual Drawing Rate	Computer Input Rate	370/168 Computer Time Used for Input	Autologic APS-4 Time Used For Output
Ring Image	Not Available	9 Str./Hr.	1.2 sec	5 sec
Structure Cuts	8 Str./Hr.	15 Str./Hr.	1.2 sec	6 sec
CA Issues	10 Str./Hr.	19 Str./Hr.	1.1 sec	4 sec

Figure 9. Manual and computer rates.

are derived from a substantial volume of information, representing over 85000 chemical structure diagrams.

In addition to being somewhat less expensive than traditional methods, computer composition offers several other advantages:

- The quality of the image is, for the most part, improved, since lines are always of a uniform width and length, bond angles are exact, and the image in never "smudged".
- Since most of the placement of lines and characters is done by computer program, there is less variation from one diagram to the next.
- Since the information can be verified by computer editing, there is less chance for errors to appear in the finished product.

Another major advantage of computer composition is that, since the information is in computer-readable form, it can be reused, each time for a fraction of what it would cost if the same information had to be manually processed. At CAS, one-time input and multiple use of textual data have been practiced for many years. We are now able to extend this technique to graphical data. For instance, only about 25% of the structure diagrams that appear in any given volume of the Chemical Substance Index are newly identified during that volume's processing. The remaining 75% are retrieved from the Structure Cut File archives. These same structure diagrams will be used again during the preparation of the Collective Chemical Substance Index, and in other publications, such as the Parent Compound Handbook. For CA issues, about 20% of the structure diagrams are retrieved from the machine-readable file of repetitive structures. The abstracts that appear in CA, along with any structure diagrams contained in the abstracts, are used multiple times, for example, in specialized publications such as *Selenium and Tellurium Abstracts* and *CA Selects*.

FUTURE WORK

Work is currently underway to replace the PDP-15 with newer hardware, consisting of a DEC PDP-11/34 and two Hughes Aircraft Corporation Conographic-9 graphics terminals. Very few changes are being made in the IBM 370 part of the Graphics System, except that the data commu-

nications subsystem is being converted from BTAM to TCAM (*B*asic *T*elecommunications *A*ccess *M*ethod; *T*elecommunications *A*ccess *M*ethod). OLSIS is being extended to handle a wider variety of character fonts and to provide a facility for handling circles and arcs. Sometime in 1978 OLSIS will be used to process the chemical structure diagrams that appear in several of the primary journals published by the American Chemical Society. A program (similar to OLSIS) is under development which will provide the capability to computer-compose the mathematical equations that appear in the primary journals. A capability for handling tables will be added sometime in 1978.

ACKNOWLEDGMENT

CAS is pleased to acknowledge the funding of this work by the National Science Foundation (Contract C656).

LITERATURE CITED

(1) E. J. Corey and W. T. Wipke, "Computer-Assisted Design of Complex Organic Syntheses", *Science*, **166**, 178–192 (1969).
(2) P. A. Castleman et al., "The Implementation of the PROPHET System", *Proc. Nat. Comput. Conf., 1974*, **43**, 457–468 (1974).
(3) R. J. Feldmann and S. R. Heller, "An Application of Interactive Graphics—The Nested Retrieval of Chemical Structures", *J. Chem. Doc.*, **12**, 48 (1972).
(4) B. E. H. Saxberg, D. S. Blom, and B. R. Kowalski, "A Flexible Interactive Graphics System for Searching Atom Connectivity Matrices", *J. Chem. Inf. Comput. Sci.*, **16**, 234–238 (1976).
(5) PDP-15 GRAPHIC-15 Reference Manual, Digital Equipment Corp., 1971.
(6) D. E. Knuth, "The Art of Computer Programming", Vol. 1, "Fundamental Algorithms", Addison-Wesley, Reading, Mass., 1968, p 306.
(7) N. A. Farmer, F. A. Tate, C. E. Watson, and G. A. Wilson, "Extension and Use of the CAS Chemical Registry System", CAS Report No. 2, April 1973, pp 3–10.
(8) P. G. Dittmar, R. E. Stobaugh, and C. E. Watson, "The Chemical Abstracts Service Chemical Registry System. I. General Design", *J. Chem. Inf. Comput. Sci.*, **16**, 111 (1976).
(9) S. J. Tauber and K. Rankin, "Valid Structure Diagrams and Chemical Gibberish", *J. Chem. Doc.*, **12**, 30 (1972).
(10) P. G. Dittmar, J. Mockus, and K. M. Couvreur, "An Algorithmic Computer Graphics Program for Generating Chemical Structure Diagrams", presented at the 169th National Meeting of the American Chemical Society, Philadelphia, Pa., April 1975; *J. Chem. Inf. Comput. Sci.*, **17**, 186 (1977).
(11) N. A. Farmer and J. C. Schehr, "A Computer-Based System for Input, Storage, and Photocomposition of Graphical Data", *Proc. ACM*, 1974, 563–570 (1974).
(12) C. Christensen and E. N. Pinson, "Multifunction Graphics for a Large Computer System", Proceedings of the 1967 FJCC, Vol. 31, AFIPS Press, Montvale, N.J., pp 697–711.

Computer-Assisted Analysis of Infrared Spectra of Nitrogen-Containing Organic Compounds

Y. MIYASHITA, S. OCHIAI, and S. SASAKI*
Miyagi University of Education, Aoba, Sendai, Japan

Received June 15, 1977

Set theory has been applied to the computer-assisted analysis of infrared spectra of the compounds $C_nH_{2n+3}N$, $C_nH_{2n+1}N$, and $C_nH_{2n+3}NO$.

A computer system to enumerate all the possible structural isomers of an unknown compound based on its molecular formula and partial structures provided by the analysis of its spectral data, NMR, IR, and so forth (system CHEMICS), has been reported.[1] However, the system works to elucidate only the structures of organic compounds with C, H, and O. In other words, the computer-assisted analysis of spectra in system CHEMICS has been designed for only compounds with C, H, and O.

To widen the applicability of the system, IR spectral analysis by computer for compounds containing N in addition to C, H, and O has been investigated by using set theory. In the present paper, the program system to analyze IR spectra of $C_nH_{2n+3}N$ (A), $C_nH_{2n+1}N$ (B), and $C_nH_{2n+3}NO$ (C) will be described.

EXPERIMENTAL

A small computer (FACOM 230-15) was used and the program was written in Fortran IV. IR spectra were measured with IR-27G (Shimadzu) and all spectra have been recorded using liquid film for pure liquid. IR spectra were also collected from IRDC cards,[2] Standard Sadtler-Infrared Prism,[3] the Aldrich Library of Infrared Spectra,[4] and some other sources.

SYSTEM

Analysis of IR Spectra. An infrared spectrum includes information on positions and intensities of bands. Characteristic group frequencies have been widely used for qualitative structural analysis of compounds because the approximate constancy of positions of group frequencies has been well established. On the other hand, the intensities of the characteristic bands have been used for the quantitative analysis of functional groups.[5] However, the intensity is rather difficult to express as a constant of the spectrum because the value varies according to the differences in experimental conditions (instrument, slit width, and others), although methods for converting apparent intensities into true molar absorption coefficients and integrated areas have been reported.[6]

To establish the effective analysis of IR spectrum for CHEMICS, we tried to use the intensities of the characteristic bands as well as the characteristic frequencies of a sample compound. The former (intensity of band) is especially useful in determining the presence of a particular partial structure. The molar extinction coefficient (ϵ) is adopted as the information of band intensity. Some of the IR spectral data of A-, B-, and C-type compounds were measured in our laboratory and others were collected from the literature.

Not all the vibrations are good group frequencies because some group frequencies overlap each other in certain compounds. Spectral regions of group frequencies which seem to be effective for the analysis are decided as described below from the IR spectra collected. Nineteen kinds of substructures, named "components" (1–19), previously defined for the analytical work and spectral region(s) corresponding to some of them (6–8, 10–19), are shown in Figure 1 (all of the components will be shown later). Three regions are defined as follows: region 1 (3500–3150 cm^{-1}), region 2 (1690–1630 cm^{-1}), and region 3 (1629–1560 cm^{-1}).

Needless to say, one is always given the apparent peak intensity which never exceeds the true peak intensity in IR spectral charts. It is well known that when the ratio of the spectral slit width of a spectrometer and the half band width ($\Delta\nu_{1/2}$) is less than 1/3, the error of the difference in intensity between the apparent peak and the true peak is not beyond 10%.

Thus $\Delta\nu_{1/2}$ of several typical compounds were measured to examine the influence of the effect of finite slit width. The observed values of $\Delta\nu_{1/2}$ of group frequencies which seem to be effective for the analysis are shown in Table I. Here, $\nu^{a}_{NH_2}$, $\nu^{s}_{NH_2}$, and $\delta^{s}_{NH_2}$ stand for the frequencies of asymmetric, symmetric stretching, and scissors modes of an amino group,