Figure 9. Similar procedures can be used for fisular ring systems of more than one ring size. In these ring systems, lowest numbering is obtained if numbering begins with the largest ring, as in Figure 10.

3. SUMMARY

Taylor's admittedly incomplete nomenclature proposal for organic ring systems has been shown to be amenable to improvement and extension. Taylor's Latin distributive prefixes and Greek number names are irregular and have been modified to parallel the mathematical system of numbering. Incorporation of graph theoretical principles into Taylor's proposal has permitted its extension to include ring systems containing atoms of valency (or connectivity) greater than 4; i.e., the modified Taylor nomenclature proposal can name any ring system, including metallocenes and "cages". However, a compatible system for naming acyclic graphs must be added before a comprehensive nomenclature system can be developed and compared with the nodal nomenclature approach of choosing a path in order to determine which is the better basis for a comprehensive graph-based nomenclature system.

REFERENCES AND NOTES

(1) Goodson, A. L. "Graph-Based Chemical Nomenclature. I", preceding paper in this issue.
(2) Taylor, F. L. "Enumerative Nomenclature for Organic Ring Systems", *Ind. Eng. Chem.* **1948**, *40*, 734-8.
(3) Terent'ev, A. T.; Potapov, V. M.; Kost, A. N.; Tsukerman, A. M. "Systematic Nomenclature of Organic Compounds", *Vestn. Mosk. Univ.* (10), No. 6, *Ser. Fiz.-Mat. Estestv. Nauk*, No. 4, **1955**, 97-134.
(4) Balaban, A. T.; Harary, F. "Chemical Graphs. V. Enumeration and Proposed Nomenclature of Benzenoid Cata-Condensed Polycyclic Aromatic Hydrocarbons", *Tetrahedron* **1968**, *24*, 2505-16.
(5) Balaban, A. T. "Chemical Graphs. VII. Proposed Nomenclature of Branched Cata-Condensed Benzenoid Polycyclic Hydrocarbons", *Tetrahedron* **1969**, *25*, 2949-56.
(6) Balaban, A. T.; Schleyer, P. v. R. "Systematic Classification and Nomenclature of Diamond Hydrocarbons. I. Graph-Theoretical Enumeration of Polymantanes", *Tetrahedron* **1978**, *34*, 3599-609.
(7) Dyson, G. M. "La Nomenclature des Hydrocarbures Polycycliques Système Dyson-Taylor-Patterson", *Bull. Soc. Chim. Fr.* **1957**, 45-52.
(8) Fletcher, J. H.; Butler, J. "Modified Taylor Nomenclature for Ring Systems", private communication.
(9) Fletcher, J. H.; Butler, J. "A Systematic Nomenclature for Organic Aliphatic Systems", private communication.
(10) Lozac'h, N.; Goodson, A. L.; Powell, W. H. "Nodal Nomenclature—General Principles", *Angew. Chem., Int. Ed. Engl.* **1979**, *18*(12), 878-99.
(11) Baeyer, A. "Systematik und Nomenclatur Bicyclischer Kohlenwasserstoffe", *Ber.* **1900**, *33*, 3771-5.
(12) Zamora, A. "An Algorithm for Finding the Smallest Set of Smallest Rings", *J. Chem. Inf. Comput. Sci.* **1976**, *16*, 40-3.
(13) Schmidt, B.; Fleischhauer, J. "A Fortran IV Program for Finding the Smallest Set of Smallest Rings of a Graph", *J. Chem. Inf. Comput. Sci.* **1978**, *18*, 204-6.
(14) "Ring Analysis Index, Index of Parent Compounds I", Parent Compound Handbook; Chemical Abstracts Service: Columbus, Ohio, Jan 1979.
(15) A system based on numbers for naming elements above 100 has been proposed by the IUPAC Commission on Nomenclature of Inorganic Chemistry. See "Recommendations for the Naming of Elements of Atomic Numbers Greater than 100", *Pure Appl. Chem.* **1979**, *51*, 381-4.
(16) Not used, but added for completeness.
(17) Used only when following a multiple of ten, as in 11, 21, 101, etc.
(18) Icosa is now recommended by IUPAC. See "Nomenclature of Organic Chemistry, Sections A, B, C, D, E, F and H"; IUPAC, Pergamon Press: Oxford, 1979; p 5.
(19) These terms are not used alone because the smallest ring has three sides.

Off-Line Input of Chemical Structures with a Low-Cost Microprocessor-Controlled Semigraphical CRT

W. KALBFLEISCH* and G. OHNACKER

Research Department, Dr. Karl Thomae GmbH, 7950 Biberach/Riss, Germany

Received September 29, 1979

Use of a microprocessor-controlled device with semigraphical CRT, light pen, and keyboard in conjunction with a direct access storage medium facilitates and speeds up the input of structural information.

In an earlier publication[1] we outlined the basic concepts of our computer-based documentation system for company research results CCBF (Computer Processing of Chemical and Biological Facts). The system has been further developed in the meantime to keep pace with progress in computer technology. Experience, gained in the daily use of the system, was evaluated and embodied in a series of system expansions and improvements, which are to be reported in the near future.[2] In addition, the steadily increasing costs of manpower forced us to look for a less time consuming means for the input of chemical and, especially, structural information.

Of the known systems, electronic devices using the RAND tablet, graphical CRT, or TV cameras[3-7] had to be excluded because of their high costs and the lack of on-line access to a corresponding computer in our area. Of the less expensive mechanical devices, even the most advanced chemical teletype[8,9] still had apparent drawbacks which we sought to overcome.

The availability of inexpensive microcomputer modules and the fact that floppy disks can serve as an off-line recording medium as well as a random access storage medium allowed the realization of a practice-oriented system tailored to user suggestions and requirements. As can be seen in Figure 1, the device consists of five functional units: (1) control unit containing a MOS Technology 6502 microprocessor, read-only-memory (ROM) program storage, random-access-memory (RAM) as display and working storage, input-output-control and videocontrol; (2) 12-inch industrial b/w video monitor; (3) light pen with integrated amplifier; (4) teletype keyboard, supplemented by a block of 12 keys, which are frequently required with special symbols (especially bond symbols); (5) dual-device floppy disk storage system (IBM 3740 format).

A general flow diagram of the processing program is given in Figure 2. It should be stressed initially that the device can work in two operating modes: "normal-mode", i.e., light pen active; "text-mode", i.e., light pen inactive. The selected mode is indicated by letters A (from German word *A*us = off) and E (from *E*in = on) behind the displayed command word TXT. After initializing, the working grid is displayed, showing three

0095-2338/80/1620-0176$01.00/0

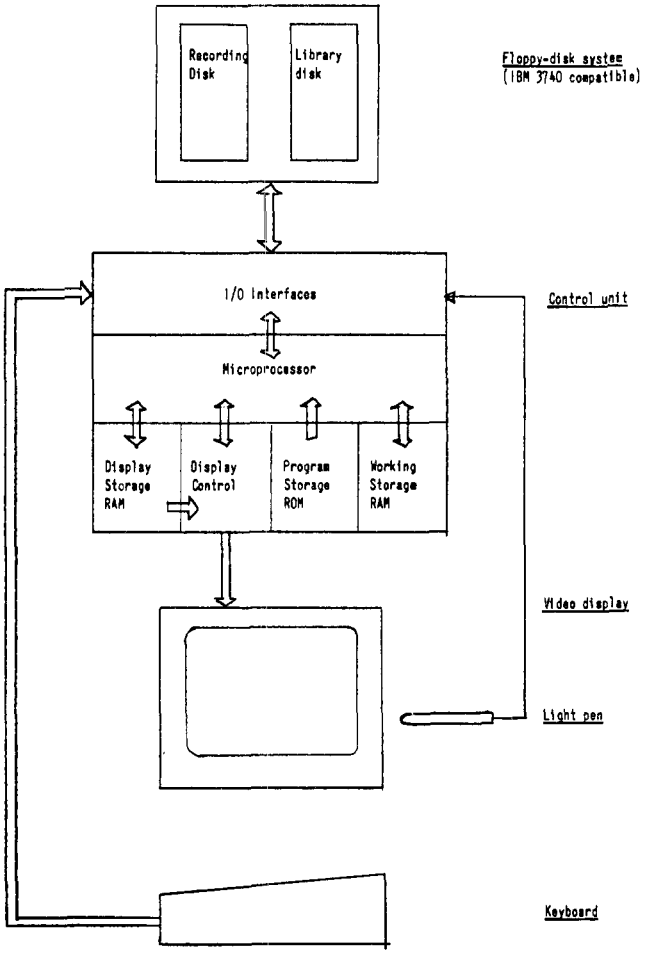

Figure 1. Functional units and their interconnections.

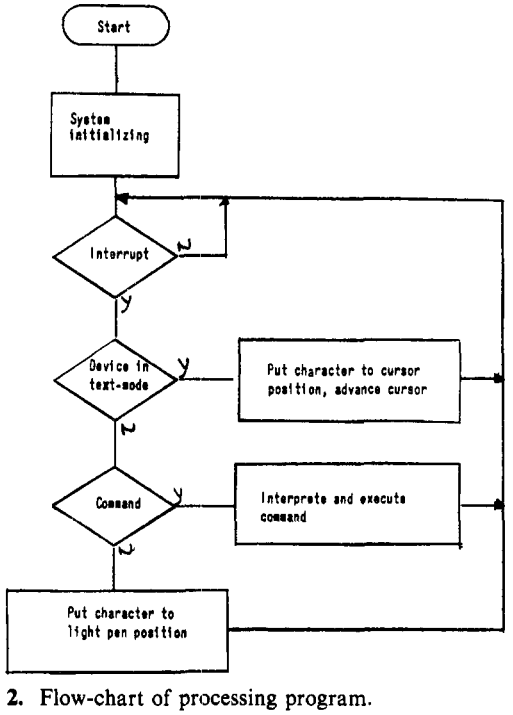

Figure 2. Flow-chart of processing program.

different parts (Figure 3): formula section with numbering of 19 rows and 47 columns, text section with 8 rows of 56 positions, and command section (directive menu).

At this point, the program loops, awaiting an interrupt signal. In "normal-mode", the interrupt is generated by co-

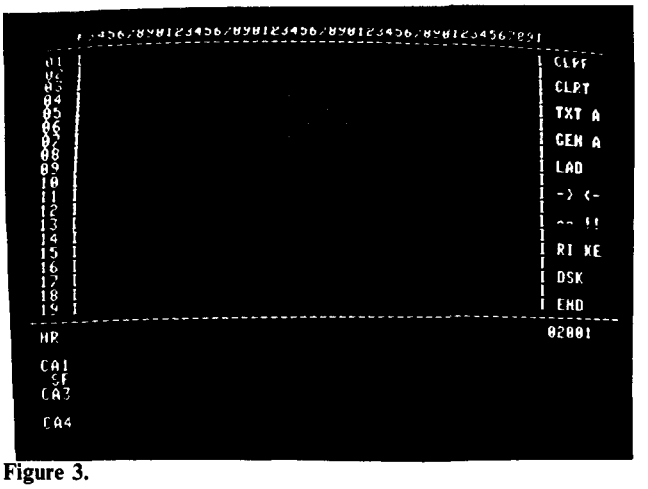

Figure 3.

inciding signals from light pen and keyboard; in "text-mode", any keyboard action is sufficient to constitute an interrupt signal. The interrupt routine first distinguishes the momentarily active mode. In "text-mode", the keyed character is put into the cursor-indicated position, the cursor is advanced, and the program returns to the waiting loop. In "normal-mode", the routine has to determine if the light pen position referred to the command section; in this case, the command is interpreted and executed. Otherwise, the keyed character is placed at the indicated light pen position and the program again returns to the waiting loop.

By this means, structure information can be "drawn" onto the videodisplay character by character. The following description of the command section will show that various options are provided, which offer time-saving short-cuts.

1. CLRF: Erases the formula section on the screen.
2. CLRT: Erases the text section on the screen.
3. TXT: The system is switched to "text-mode", the light pen is deactivated, and the cursor is brought to the first position of the text section, normally the input position for the serial number. Text-mode can also alternatively be entered by bringing the "cursor" to any position of the formula or text section via the light pen. In this manner it is possible to write comments or names in the formula part and/or correct or edit previously written text in both sections. In the "text-mode", the keys "space" (→), "rub-out" (←), "line-feed" (↓), and "arrow-up" (↑) are used for cursor control "right", "left", "down", and "up", respectively. The "return" key moves the cursor to the first position of a new line; key "alt-mode" is used to leave "text-mode". All other keys are decoded normally and displayed.
4. GEN: This function simulates "normal" light pen application of graphical systems and allows automatic generation of single-bond symbols between the last two indicated light pen positions. "GEN" function is mainly used in drawing expanded ring systems, for instance, bicyclic or tricyclic rings. An error signal and a function reset are generated if the positions are not horizontally, vertically, or diagonally connectable. Normally the function is left by hitting the "alt-mode" key.
5. With commands → ← ↑ ! the information of the formula section can be shifted one column to the right or left, or one row up or down.
6. RI: With this function, a series of differently positioned five-, six-, and seven-membered ring skeletons can be generated. To accomplish this, the appropriate code is brought into the desired position and then the light pen is pointed to command RI. All skeletons are generated with carbon-atom and single-bond symbols; multiple-bond symbols or heteroatoms can be applied by subsequent overwriting. Condensed and spiro-connected rings can easily be constructed by repeated

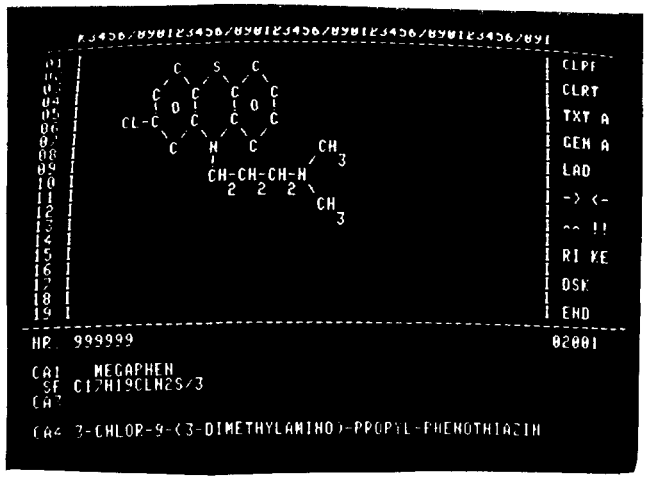

Figure 4.

use of the RI functions. Drawing of a steroid skeleton, for instance, requires only eight keystrokes.

7. KE: Generation of hydrocarbon chains is released by this command. A code from $n = 1$ to 9 displays n-membered chains of the form CH_3–CH_2 ..., while the letters A to I generate the corresponding condensed forms C_nH_{2n+1}. As above, the generated information can be easily modified (multiple bonds, heteroatoms, branches).

The following three commands release diskette operations. It has to be pointed out initially that the left disk-drive serves as the recording medium for the structures to be further processed in the documentation system, while the drive on the right side serves as a direct access medium to a user-constructed library of prerecorded structures.

8. END: This command stores the contents of the display on the recording disk. Only the nonblank characters from the formula section are stored along with their coordinates. The next free disk address (track and section address) is displayed on the first line of the text section. The variable amount of disk space is dependent on the size of the formula and on the amount of text the recorded information requires. With an average of three sectors per document one diskette can hold about 700 documents. The recorded diskettes serve as the input medium for our weekly batch-mode system update.

9. DSK: This command provides a series of diskette operations. Depending on which key is depressed, the following subcommands are executed:

- O *O*pen disk, reset to track 2 sector 1
- C *C*lose disk, write end-of-file record
- M *M*ove contents of left disk to right disk
- S *S*ave the displayed structure on the library (right) disk; the catalog index number must be provided in field NR of the text section
- D *D*elete the cataloged structure found in field NR of the text section

10. LAD: Fetch and display the cataloged structure found in field NR of the text section.

During input, the program checks for adherence to some simple syntactical rules, for instance, validity of codes and subcommands, presence of a serial number, and so on. In case of error, the type of error is indicated both optically (on the display) and acoustically (a buzzer signal).

CONCLUSION

The device has been in operation for more than 2 years during which time several thousand structures have been processed. As expected, it exhibited the special advantage that the displayed formula and text remain accessible even after storing on the recording disk. Thus a new structure can be created by overcoding, appending, or partially clearing the previously drawn structure. It is obvious that this feature is particularly advantageous in research documentation, especially in pharmaceutical research, where frequently only minor structural variations have to be handled during a particular input session.

High importance must also be given to the possibility of direct access to cataloged structures. Basic structures, to be expected from the future research program, can be stored ready for instant retrieval when required. The amount of time needed to produce a single document is naturally dependent on the complexity of the structure, the amount of text, and the utilization of the system capabilities. On the average, a turnover of about 30 documents per hour can be expected, the ratio between the expenditure for structure and text input being about 40:60.

The document shown in Figure 4, for instance, requires 25 keystrokes to produce the structure and 77 keystrokes for text information (a detailed description of the systems coding conventions, the possibility of hydrogen suppression, use of abbreviations, and so on, are given in ref 1 and 2).

Inspection of figures given by authors[6,9] shows that our device compares very favorably with known off-line systems. With a price of about $10000, the device is also competitive with the various chemical typewriters as well as with video displays utilizing thumbwheel- or key-controlled cursor positioning. The device has been developed and constructed on a noncommercial basis. The authors will be pleased to provide further information.

LITERATURE CITED

(1) Ohnacker, G.; Kalbfleisch, W. *Angew. Chem., Int. Ed. Engl.* **1970**, *9*, 605.
(2) Becker, J.; Jung, D.; Ohnacker, G.; Kalbfleisch, W., submitted for publication.
(3) Woodward, W. S.; Isenhour, T. L. *Anal. Chem.* **1974**, *46*, 422.
(4) Farrell, C. D.; Chauvenet, A. R.; Koniver, D. A. *J. Chem. Doc.* **1971**, *11*, 52.
(5) Feldman, R. J.; Heller, S. R.; Shapiro, K. P. *J. Chem. Doc.* **1972**, *12*, 41.
(6) Brugger, W. E.; Jurs, P. C. *Anal. Chem.* **1975**, *47*, 781.
(7) Howe, W. J.; Hagadone, J. R. "Retrieval of Medicinal Information", ACS Symposium Series No. 84; American Chemical Society: Washington, D.C., 1978.
(8) Mullen, J. M. *J. Chem. Doc.* **1967**, *7*, 88.
(9) Feldman, A. *J. Chem. Doc.* **1973**, *13*, 53.