(22) Norris, J. R.; Warden, J. T. A Rapid Response 100 kHz ESR Spectrometer. *EPR Lett.* **1980**, *1*, 2-3.
(23) PHYS(2) in *Berkeley Unix Programmer's Manual: PDP-11 Version 2.9*, 7th ed. (Second Berkeley Software Distribution); University of California at Berkeley: Berkeley, CA, 1983; Vol. 2.
(24) Ferrin, Thomas. RTP(2) in *Berkeley Unix Programmer's Manual: PDP-11 Version 2.9*, 7th ed. (Second Berkeley Software Distribution); University of California at Berkeley: Berkeley, CA, 1983; Vol. 2.
(25) LD(1) in *Berkeley Unix Programmer's Manual: PDP-11 Version 2.9*, 7th ed. (Second Berkeley Software Distribution); University of California at Berkeley: Berkeley, CA, 1983; Vol. 1.
(26) Palace, Gerard; Franke, James E.; Warden, Joseph T. Is phylloquinone an obligate electron carrier in photosystem I? *FEBS Lett.* **1987**, *215*, 58-62.

CHEMLEARN: Microcomputer-Based Training for CHEMLINE, an Alternative to Formal Classroom Training†

MELVIN L. SPANN* and MIRIAM L. PERKINS
National Library of Medicine, Bethesda, Maryland 20894

Received September 8, 1989

CHEMLINE (CHEMical Dictionary OnLINE), the National Library of Medicine's online chemical dictionary file, is primarily used by librarians and information scientists to enhance the retrieval of bibliographic information associated with chemical substances. This paper will discuss CHEMLEARN, a microcomputer-based training program designed to provide inexpensive, easily accessible self-instruction as an alternative to formal classroom training in the use of CHEMLINE. CHEMLEARN allows novice users to learn at their own pace and with considerable program feedback. In addition, it provides the skilled searcher with a way to reinforce or recall previously learned search techniques without incurring online charges. CHEMLEARN may be used in place of formal training or as a precursor to or a refresher following formal training.

INTRODUCTION

CHEMLINE is the National Library of Medicine's (NLM) online, interactive chemical dictionary file created by NLM's Toxicology Information Program in conjunction with Chemical Abstracts Service (CAS).[1] It serves as an adjunct file for the search and retrieval of bibliographic and factual information pertaining to chemical substances. In fact, CHEMLINE can be viewed as the focal point for chemical searching of the Library's online retrieval services. The file provides a mechanism whereby information on over 900 000 chemical substances can be searched and retrieved online. It contains CAS Registry Numbers, molecular formulas, CAS chemical index nomenclature, generic and trivial names, classification codes, and a locator designation that points to other files in the NLM system and the Toxic Substances Control Act (TSCA) Inventory. In addition, where applicable, each record contains ring information including number of component rings within a ring system, ring sizes, ring elemental compositions, and component line formulas. This permits the searcher to refine and amplify chemical search terms in searching the selected databases.

The majority of CHEMLINE users are search intermediaries, librarians or information scientists conducting searches for scientists or health professionals.[2] NLM has been training search analysts since the implementation of its online computer system, MEDLARS (Medical Literature Analysis and Retrieval System), in the mid 1960s.[3] Currently, NLM offers two formal training courses geared toward the search intermediary which include segments on CHEMLINE. They are a 3-5-day modular introductory course titled "The Fundamentals of MEDLARS Searching" and a week-long modular sequel course titled "The Follow-Up to the Fundamentals of MEDLARS Searching". As useful as these courses are, they have the drawback of requiring the search analyst to travel to a training site for a scheduled course which may be offered in their locale only a few times per year and which can be both costly and time-consuming. In addition, these courses do not permit the degree of self-paced instruction that some learners desire or need, especially considering the highly technical nature of CHEMLINE. Furthermore, due to recent increases in the number of users, the capacity to train these individuals in a timely manner is being fully taxed.

A main-frame computer-assisted instruction program for CHEMLINE was developed at the National Library of Medicine in 1979[4] to meet the needs of searchers who wanted to learn basic concepts involved in searching CHEMLINE in the privacy and convenience of their own offices. It was designed to be accessed through the then predominant hard-copy terminal connected via modem to the main-frame computer at NLM. With the increased availability of microcomputers, it became clear that yet another method of instructing searchers in the use of CHEMLINE would be perhaps more practical. Microcomputer-based training programs provide economical, convenient self-instruction for database searchers and allow for use of enhanced visual and graphic capabilities in conveying technical information. Therefore, in the spring of 1988 NLM released CHEMLEARN, a microcomputer-based training program, to address the needs of CHEMLINE search analysts who want to learn or continue learning at their own pace, on their own equipment, in their home or office, and at an economical rate.[5] After over a year of user feedback, CHEMLEARN, version 2.0, was released incorporating many new capabilities and techniques.

A technical database such as CHEMLINE is only worthwhile if the audiences for whom it is intended are using it. As with any technology or tool, some training is necessary for searchers to use it effectively. NLM has developed a variety of learning methodologies including formal classes, printed manuals, and computer-based training to address this need. One of the few well-established principles of educational psychology is that people learn via a variety of styles, rates, and sequences. Whereas some people learn most efficiently in a group-paced classroom environment, others benefit more from a self-paced, highly structured computer-based environment. At the same time, temporal and financial logistics may prohibit some potential users from availing themselves of in-person classroom instruction, and, therefore, independent

† Presented at the 198th National Meeting of the American Chemical Society, Miami Beach, FL, Sept 1989, paper CINF 30.

learning opportunities must be provided to them. As a result, the provision of microcomputer-based tutorials such as CHEMLEARN is probably best regarded as one alternative type of learning opportunity in an "arsenal" of learning opportunities provided by NLM.[6]

CONTENT

CHEMLEARN is organized into seven chapters. The introduction is presented in Chapter 1. Chapters 2, 4, and 6 provide interactive instruction on the content and use of the CHEMLINE data elements. Chapter 2 covers the basic data elements used when CHEMLINE is searched. The name and formula fragment data elements are covered in Chapter 4 and the ring data elements in Chapter 6. Chapters 3, 5, and 7 provide simulated searches that illustrate actual online sessions. The following diagram illustrates the three levels of complexity (i.e., basic, advanced, and specialized), each containing an instruction chapter followed by a search chapter, which make up the CHEMLEARN program.

1. Introduction

2. Basic Data Elements 3. Simulated Searches

4. Fragment Data Elements 5. Simulated Searches

6. Ring Data Elements 7. Simulated Searches

A hierarchical listing of the sections in each of the CHEMLEARN chapters is

- Introduction
- Basic Data Elements
  - Registry Number
    - Registry Number
    - Other Registry Number
    - Related Registry Number
    - Searching By Registry Number
  - Molecular Formula
    - Molecular Formula
    - Searching By Molecular Formula
    - Molecular Formula Truncation
  - CAS Type 1 Name
    - CAS Type 1 Name
    - Searching By Chemical Name
    - Chemical Name Truncation
  - Synonym
    - Synonym
    - Searching By Common Name
    - Searching By Chemical Name
    - Common Name Truncation
  - Locator
    - Locator
    - Searching By Locator
  - Classification Code
    - Classification Code
    - Searching By Classification Code
- Simulated Searches With Basic Data Elements
  - Synonym/Related Registry Number Search
  - Synonym/Molecular Formula Search
  - Registry Number/CAS Type 1 Name Search
  - Classification Code/Locator Search
- Fragment Data Elements
  - Name Fragment
    - Name Fragment
    - Searching By Name Fragment
    - Name Fragment Truncation
  - Formula Fragment
    - Formula Fragment
    - Searching By Formula Fragment
- Simulated Searches With Fragment Data Elements
  - Name Fragment Search
  - Name Fragment/Formula Fragment Search
- Ring Data Elements
  - Number of Rings
  - Ring Size
  - Ring Elemental Analysis
  - Component Line Formula
  - Searching By Ring Data Elements
- Simulated Searches With Ring Data Elements
  - Ring/Name Fragment Search
  - Ring/Name Fragment/Formula Fragment Search

The information in the CHEMLEARN chapters is structured to present the simpler concepts and skills first and progressively evolve to the more difficult rules and concepts. As a user proceeds through the program, a knowledge base is being developed in a spiral manner. That is, a user is exposed to the same concepts at more than one point; however, each successive exposure is at a higher cognitive level.

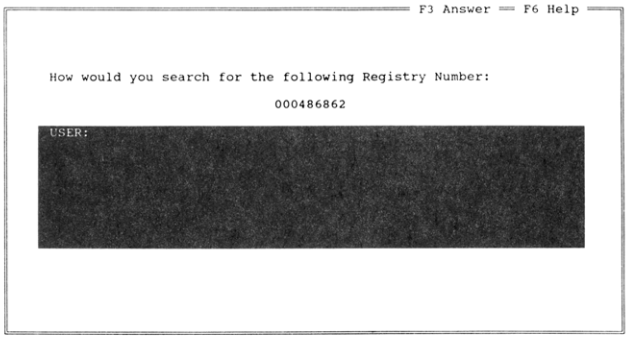

Figure 1. Registry Number completion problem.

INSTRUCTIONAL DESIGN

Perhaps the most important design aspect of CHEMLEARN is its systematic practice and diagnostic feedback. This is especially important considering the highly technical nature of the CHEMLINE database and the fact that the majority of the users are search intermediates. Users interact with CHEMLEARN as they would with CHEMLINE itself, entering a letter, word, or phrase at the **USER**: prompt. CHEMLEARN responses are displayed following the **PROG**: prompt. In the instruction chapters, CHEMLEARN provides user-specific feedback for correct and incorrect answers to both completion and multiple-choice problems. When the correct answer is entered, the program reinforces or summarizes the concepts essential to the problem. When errors are made, the program displays diagnostic feedback and prompts the user to try again or have the correct answer provided. For completion problems the program matches the user's entry to a list of anticipated errors that were generated on the basis of observing trainees construct searches with similar problems in classroom settings. Thus, problems and corresponding matching algorithms are designed to detect common errors derived from previous training experience. Figures 1 and 2 show examples of completion problems from the Basis Data Elements chapter illustrating a Registry Number search and a chemical name search, respectively. Tables I and II give the corresponding program responses for correct answers and anticipated errors. Figure 3 shows a sample multiple-choice problem from the Fragment Data Elements chapter illustrating a name fragment search. The corresponding program responses for the correct and incorrect answers are given in Table III. The error-diagnosing methods used in CHEMLEARN simulate the learner misconception model used in Intelligent Computer-Assisted Instruction (ICAI) programs.[7]

In the simulated search chapters, CHEMLEARN detects incorrect search statements or commands as well as appropriate responses that deviate from program directions and displays

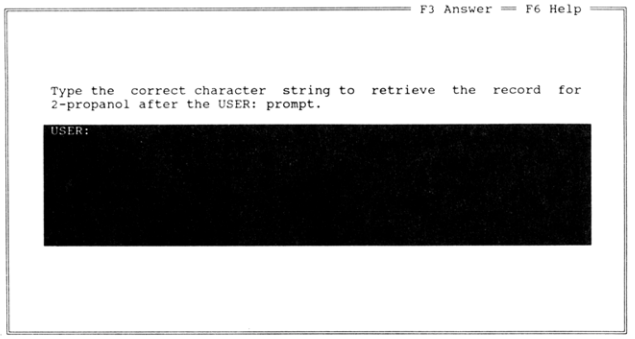

Figure 2. Chemical name completion problem.

Table I. Program Responses for Correct Answers and Anticipated Errors for the Registry Number Problem in Figure 1

user entry	program response
correct answer	Very good! Registry Numbers are searched and displayed in a hyphenated format with leading zeros dropped.
error 1 (retained leading zeros)	That is not right. Remember to drop leading zeros.
error 2 (did not use hyphens)	That is not right. Remember, Registry Numbers are searched and displayed in a hyphenated format.
error 3 (hyphens misplaced)	That is incorrect. The correct hyphenated format is, from right to left, 1 digit, hyphen, 2 digits, hyphen, up to 6 digits.

Table II. Program Responses for Correct Answers and Anticipated Errors for the Chemical Name Problem in Figure 2

user entry	program response
correct answer	Right! Chemical names in the N1 field are searched directly by entering the exact character string which designates the chemical, including spaces and special characters.
error 1 (no hyphen after locant 2)	Your answer is not right. Remember to include the hyphen after the locant when entering the chemical name.
error 2 (no locant 2)	Your answer is not right. Remember to include the locant 2 when entering the chemical name.

Table III. Program Responses for Correct Answers and Incorrect Answers for the Name Fragment Problem in Figure 3

user entry	program response
A (correct for SS 1 or SS 2)	This is correct for search statement 1 or 2. However, for search statement 3 and above, single-number name fragments must be postqualified by (NF). Otherwise, the system will assume that such numbers refer to previous search statements.
B (fragments not formed correctly)	That is incorrect. According to the rules for name fragment generation, hyphens are converted to blanks and each uniquely occurring character string is saved as a name fragment.
C (correct answer)	Excellent! You chose the answer with the individual name fragments logically ANDed together and prequalified by the (NF) mnemonic.
D (did not use 2 (NF))	The record you are searching for will be included in your retrieval; however, you could further limit your retrieval by including the numerical name fragment 2 (NF).
E (did not use (NF) qualifier)	That is incorrect. Remember, you must use a mnemonic to search the NF field in CHEMLINE.

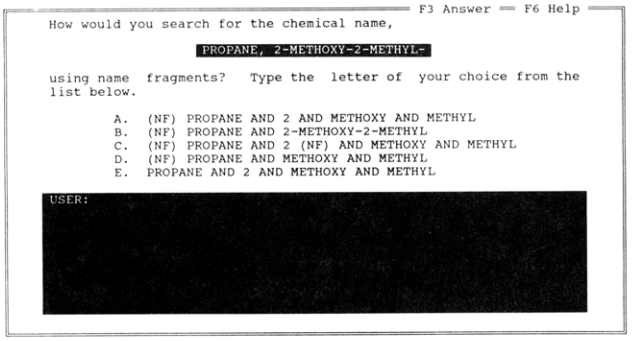

Figure 3. Name fragment multiple choice problem.

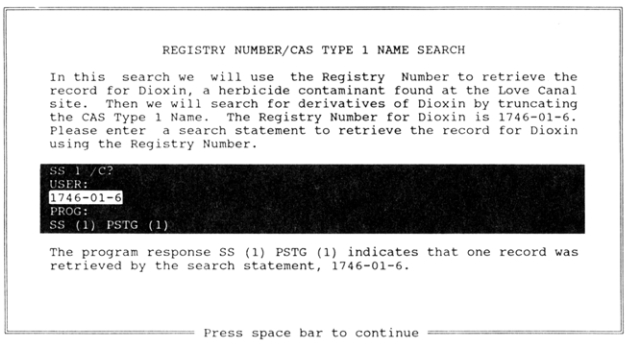

Figure 4. Simulated search interaction for dioxin.

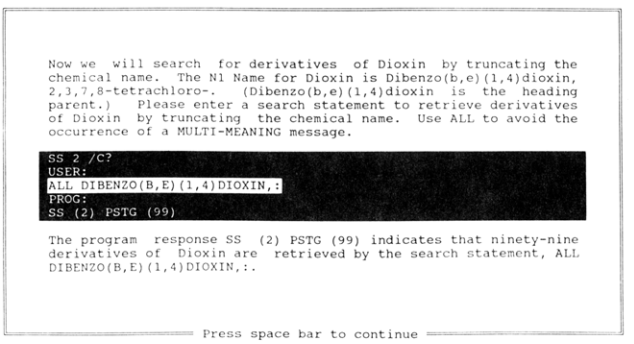

Figure 5. Simulated search interaction for dioxin derivatives.

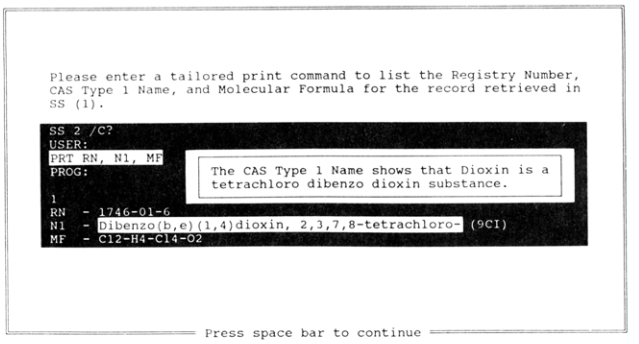

Figure 6. Print interaction from the simulated search for dioxin.

Table IV. Program Responses for Partially Correct and Incorrect Search Statements for the Simulated Search for Dioxin in Figure 4

user entry	program response
error 1 (used common name)	You entered a search statement that would retrieve the record for dioxin using the common name. For purposes of this simulation, please enter a search statement to retrieve the record for dioxin using the Registry Number.
incorrect search statements	Please enter a search statement to retrieve the record for dioxin using the Registry Number.

Table V. Program Responses for Partially Correct and Incorrect Search Statements from the Simulated Search for Dioxin Derivatives in Figure 5

user entry	program response
error 1 (did not use ALL)	You have entered a search statement that would retrieve derivatives of dioxin; however, it would generate a multimeaning message. For purposes of this simulation, please reenter your search statement using ALL.
error 2 (missing comma)	You have entered a search statement that would retrieve derivatives of dioxin; however, it would also retrieve substances with a chemical name that begins with dibenzo(B,E)(1,4)dioxin. For purposes of this simulation, please reenter your search statement with a comma immediately after the parent and before the colon.
error 3 (missing comma and ALL)	You have entered a search statement that would retrieve derivatives of dioxin; however, it would also retrieve substances with a chemical name that begins with dibenzo(B,E)(1,4)dioxin and it would generate a multimeaning message. For purposes of this simulation, please reenter your search statement with a comma immediately after the parent and before the colon and use ALL.
incorrect search statements	Please enter a search statement to retrieve derivatives of dioxin by truncating the chemical name.

Table VI. Program Response for Incorrect User Entries for the Print Interaction from the Simulated Search for Dioxin in Figure 6

user entry	program response
incorrect search statements	Please enter a tailored print command to list the RN, N1, and MF for the record retrieved in SS (1).

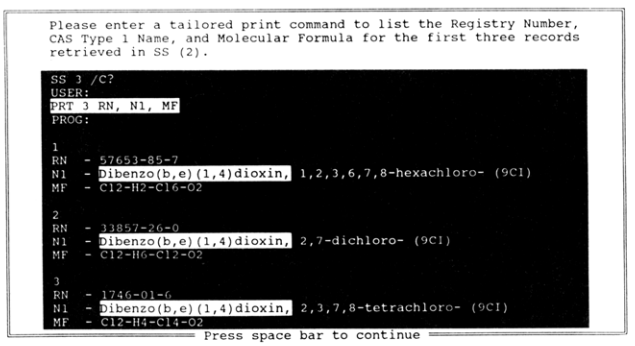

Figure 7. Print interaction from the simulated search for dioxin derivatives.

a window containing diagnostic feedback. The user is then given the opportunity to reformulate the search strategy or have the correct search strategy entered automatically. Figures 4 and 5 illustrate simulated search interactions for dioxin and dioxin derivatives, respectively. The corresponding program responses for partially correct search statements and for incorrect search statements are given in Tables IV and V.

Figures 6 and 7 show the print interactions from the simulated searches for dioxin and dioxin derivatives. The corresponding program responses for partially correct and incorrect user entries are given in Tables VI and VII. Notice, for purposes of illustration, user entries as well as search terms appear highlighted. Also, notice the window in Figure 6 providing additional information on the CAS type 1 name in the dioxin record. Similar windows are found throughout the simulated search chapters.

Optional information is provided throughout CHEMLEARN. Optional HELPS, windows summarizing the information

Table VII. Program Responses for Partially Correct and Incorrect User Entries for the Print Interaction from the Simulated Search for Dioxin Derivatives in Figure 7

user entry	program response
error 1 (all 99 records)	You have entered a print command that would list the RN, N1, MF for all 99 records retrieved in SS (2). For purposes of this simulation, please enter a tailored print command to list the RN, N1, and MF for the first three records retrieved in SS (2).
incorrect search statements	Please enter a tailored print comman to list the RN, N1, and MF for the first three records retrieved in SS (1).

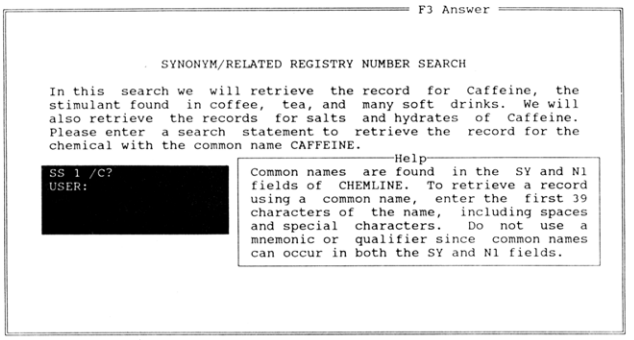

Figure 8. HELP window from the simulated search for caffeine.

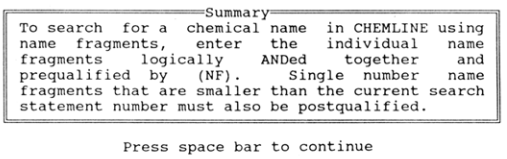

Figure 9. SUMMARY from the searching by name fragments section.

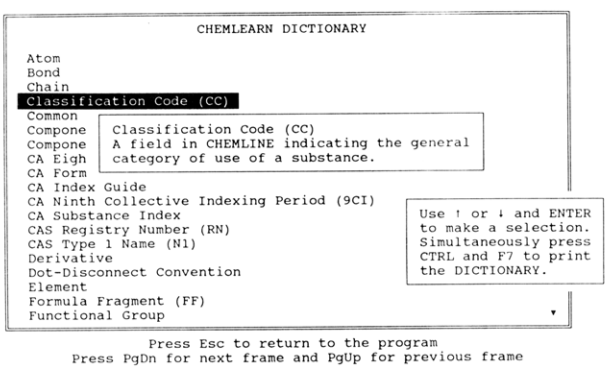

Figure 10. DICTIONARY frame displaying the definition for the classification code.

needed by the user to formulate the correct response, are available at every point of interaction in both the instruction and simulated search chapters. Figure 8 shows the HELP associated with a simulated search for caffeine. Optional SUMMARIES are available for every section in the instruction chapters. They are accessible from any point within the section and provide an excellent means of review. Additionally, the user may print all of the SUMMARIES included in a particular chapter from any point within the chapter. The SUMMARY for the Searching By Name Fragments section is shown in Figure 9. CHEMLEARN also contains a printable DICTIONARY that enables users to quickly look up definitions for data elements, basic chemical terms, standard conventions, acronyms, and publications referred to in the program. Figure 10 shows a frame from the DICTIONARY displaying the definition for the Classification Code.

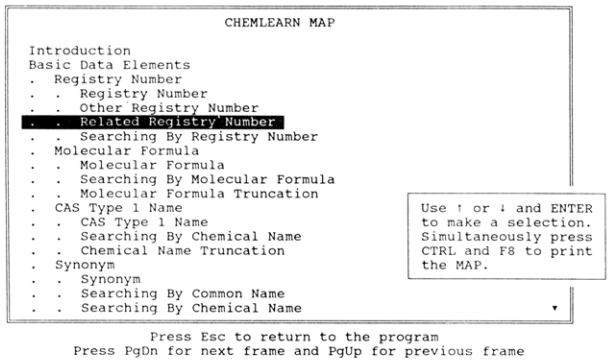

Figure 11. MAP frame.

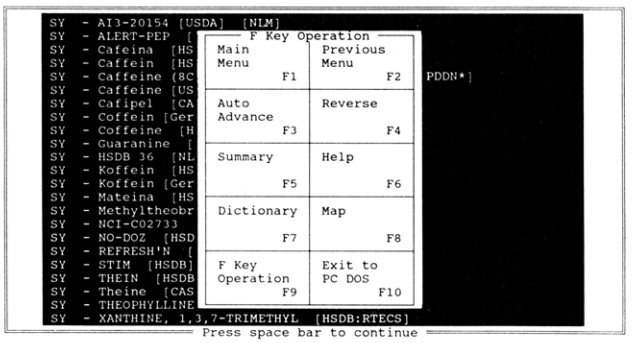

Figure 12. Function-key window.

Menus are used to access chapters and sections within chapters. A printable MAP provides a hierarchical listing of each of the CHEMLEARN sections. It serves as both an index and an overall menu. A frame from the MAP is shown in Figure 11. A tracking function monitors users' progress through the program. As each section is completed, its corresponding menu and MAP entries are highlighted. Records are maintained on the progress of the 12 most recent users. This information is used in subsequent sessions to allow these users the option of returning to the point in the program from which they exited and to indicate the topics previously completed.

Movement among menus and access to the MAP and optional information, including HELPS, SUMMARIES, and the DICTIONARY, as well as paging forward and backward through consecutive displays, automatically displaying user input, and exiting to DOS are accomplished by pressing function keys. Additionally, a window detailing each function key's operation may be displayed at any point in the program by pressing a function key. This is illustrated in Figure 12. CHEMLEARN maximizes learner control of instruction by incorporating information aids, retrieval devices, and a friendly interface. This strategy allows users to quickly access the particular portions of the program related to their individual search requirements and has been shown to be particularly effective with sophisticated learners.[8,9]

IMPLEMENTATION

CHEMLEARN was developed by using the NLM LEARN Programmer, a code generator that operates on specially formatted ASCII script files to produce PILOTplus code. PILOTplus is a high-level authoring language designed to assist the professional educator in developing interactive, microcomputer-based courseware. It is based on PILOT (Programmed Inquiry Learning or Teaching), which was developed by John Starkweather at the University of California at San Francisco in 1969.[10]

The program resides on two 5-$^1/_4$-in. low-density floppy diskettes or one 3-$^1/_2$-in. low-density disk and requires 540K RAM and DOS 2.0 or higher. It may be loaded onto a hard disk or executed directly from the floppy diskettes. It runs on IBM PCs, PC-XTs, PC-ATs, PS/2s, and compatibles and supports most monochrome, color-graphics, enhanced color-graphics, and visual color-graphics monitors. While no extra boards or external peripherals are required, a printer is desirable.

CONCLUSIONS

The purpose of this work was to provide an alternative to formal classroom training for CHEMLINE. CHEMLEARN provides the search analyst with inexpensive, easily accessible self-instruction in the use of CHEMLINE, thereby eliminating the travel, cost, and time constraints associated with formal classroom training. CHEMLEARN is the first in a series of microcomputer-based training programs for selected databases at the National Library of Medicine. Already it has served as a model for developing similiar microcomputer-based training programs for MEDLINE[11] and TOXLINE,[12] the Library's online bibliographic databases covering the biomedical, toxicological, and environmental literature. Currently, the Instructional Technology Division of the University of Georgia is evaluating the effectiveness of CHEMLEARN as a training tool for CHEMLINE. Undoubtedly, as CHEMLEARN was developed to take advantage of microcomputers, other alternatives to formal classroom training for CHEMLINE will emerge as technology continues to evolve.

ACKNOWLEDGMENT

We acknowledge the excellent programming support provided by Jeffrey Howard and thank Patricia Williams for her editorial assistance in preparing this paper.

REFERENCES AND NOTES

(1) Schultheisz, R. J.; Walker, D. F.; Kannan, K. L. Design and Implementation of an Online Chemical Dictonary (CHEMLINE). *J. Am. Soc. Inf. Sci.* **1978**, *29*, 173–39.
(2) Woelfel, J. C. *An Evaluation of CHEMLINE: The Results of a Survey of Users By the National Library of Medicine*; Market Dynamics: Vienna, VA, 1985.
(3) Kassebaum, L.; Leiter, J. Training and Continuing Education for On-Line Searching. *Med. Inf.* **1978**, *3*(3), 165–175.
(4) Spann, M. L. *CHEMLEARN*: A Computer-Assisted Instruction (CAI) Program for Online Chemical Information Searching. *Diss. Abstr. Int. A.* **1979**, *40*(10-A), 5229 (Order No. 80-8347).
(5) Perkins, M. L.; Spann, M. L.; Geasler, M. M. Introducing CHEMLEARN: Microcomputer-Based Training for CHEMLINE. *Proceedings of the 29th International Association for the Development of Computer-Based Instructional Systems, Nov 14–17, 1987.*
(6) Perkins, M. L.; Spann, M. L.; Reeves, T. C. Microcomputer-Based Training for MEDLARS Databases. *Proceedings of the 31st International Association for the Development of Computer-Based Instructional Systems, Nov 1989* (in press).
(7) Burton, R. R. Diagnosing Bugs in a Simple Procedural Skill. In *Intelligent Tutoring Systems*; Sleeman, D., Brown, J., Eds.; Academic Press: New York, 1982.
(8) Snow, R.; Lohman, D. Toward a Theory of Cognitive Aptitude for Learning from Instruction. *J. Educ. Psychol.* **1984**, *72*(3), 347–376.
(9) Tenneyson, R.; Christensen, D.; Park, S. The Minnesota Adaptive Instructional System: An Intelligent CBI System. *J. Comput.-Based Instr.* **1984**, *11*(1), 2–13.
(10) Starkweather, J. A common language for a variety of conversational programming needs. In *Readings in Computer-Assisted Instruction*; Wilson, H. A., Atkinson, R. C., Eds.; Wiley: New York, 1969.
(11) Perkins, M. L.; Spann, M. L.; Buchan, P. C. MEDTUTOR: A Microcomputer-Based Training Program for MEDLINE. *Bull. Med. Libr. Assoc.* **1989**, *77*(2), 201–204.
(12) Perkins, M. L.; Spann, M. L. TOXLEARN: Microcomputer-Based Training for TOXLINE. *NLM Tech. Bull.* **1989**, *242*, 9–14.