r_n number of rings or cycles having n vertices

REFERENCES AND NOTES

(1) Dias, J. R. *J. Chem. Inf. Comput. Sci.* **1982**, *22*, 15.
(2) Dias, J. R. *J. Chem. Inf. Comput. Sci.* **1982**, *22*, 139.
(3) Dias, J. R. *MATCH* **1983**, *14*, 83.
(4) Lewis, I. C.; Petro, B. A. *J. Polym. Sci.* **1976**, *14*, 1975.
(5) Rakauskas, R.; Bolotin, A. B. *Liet. Fiz. Rikinys* **1962**, *2*, 41.
(6) Goodwin, T. H.; Morton-Blake, D. A. *Theor. Chim. Acta* **1964**, *2*, 75.
(7) Ege, G.; Volger, H. *Theor. Chim. Acta* **1974**, *35*, 189.
(8) Tinnemans, A. H. A.; Laarhoven, W. H. *J. Am. Chem. Soc.* **1974**, *96*, 4611 and references cited therein.
(9) Zander, M., *Z. Naturforsch., A* **1975**, *30A*, 1096.
(10) Niakasuji, K.; Kazuhiro, N. *Bull. Chem. Soc. Jpn.* **1970**, *43*, 3567.
(11) Nakasuji, K.; Akujama, S.; Nakagawa, M. *Bull. Chem. Soc. Jpn.* **1972**, *45*, 875 and references cited therein.
(12) Balaban, A. T.; Harary, F. *Tetrahedron* **1968**, *24*, 2505.
(13) Knop, J. V.; Szymanski, K.; Jericevic, Z.; Trinajstic, N. *J. Comput. Chem.* **1983**, *4* (1), 23.
(14) Martin, R. H.; Eyndels, C.; Defay, N. *Tetrahedron* **1974**, *30*, 3339.
(15) Martin, R. H.; Baes, M. *Tetrahedron* **1975**, *31*, 2135.
(16) Clar, E.; Robertson, J. M.; Schlögl, R.; Schmidt, W. *J. Am. Chem. Soc.* **1981**, *103*, 1320.
(17) Jenny, W.; Peter, R. *Angew. Chem., Int. Ed. Engl.* **1965**, *4*, 979.
(18) Diederich, F.; Staab, H. A. *Angew. Chem., Int. Ed. Engl.* **1978**, *17*, (5), 372.
(19) Ege, G.; Volger, H. *Theor. Chim. Acta* **1972**, *26*, 55; *Z. Naturforsch., B: Anorg. Chem., Org. Chem.* **1972**, *27B*, 918; *Theor. Chim. Acta* **1974**, *35*, 667.
(20) Volger, H. *Croat. Chim. Acta* **1980**, *53*, 667.
(21) Agranat, I.; Hess, B.; Schaad, L. *Pure Appl. Chem.* **1980**, *52* (6), 1399.
(22) Thulin, B.; Wennerstrom, O. *Acta Chem. Scand., Ser. B* **1976**, *30B*, 369.
(23) (a) Barth, W. E.; Lawton, R. G. *J. Am. Chem. Soc.* **1971**, *93*, 1730; Hanson, J. C.; Nordman, C. *Acta Crystallogr., Sect. B* **1976**, *32B*, 1147. (b) Yamamoto, K.; Harada, T.; Nakazaki, M.; Naka, T.; Kai, Y.; Harada, S.; Kasai, N. *J. Am. Chem. Soc.* **1983**, *105*, 7171.
(24) Staab, H. A.; Graf, F. *Chem. Ber.* **1970**, *103*, 1107; Mitchell, R.; Sondheimer, F. *Tetrahedron* **1970**, *26*, 2141.
(25) For example, see Akiyama, S.; Takeuchi, Y.; Yasuhara, A.; Nishimoto, K.; Nakagawa, M. *Bull. Chem. Soc. Jpn.* **1973**, *46*, 2830 and references cited therein.
(26) Schonberg, A.; Brosowski, K.; Singer, E. *Chem. Ber.* **1962**, *95*, 1910.
(27) Wynder, E.; Hoffman, D. *Dtsch. Med. Wochenschr.* **1963**, *88*, 623.
(28) Bochar, D.; Gal'pern, E.; Gambaryan, N. *Izv. Akad. Nauk. SSSR, Ser. Khim.* **1970**, *No. 2*, 435.
(29) Laarhoven, W. T.; Veldhuis, R. G. M. *Tetrahedron* **1972**, *28*, 1823.
(30) Tochtermann, W. *Angew. Chem., Int. Ed. Engl.* **1963**, *2*, 265.
(31) Jensen, E.; Kovats, E.; Eschenmoser, A.; Heilbronner, E. *Helv. Chim. Acta* **1956**, *39*, 1051; Mayer, J.; Sondheimer, F. *J. Am. Chem. Soc.* **1966**, *88*, 603; Muth, C.; DeMatte, M.; Urbanik, A.; Isner, W. *J. Org. Chem.* **1966**, *31* 3013; O'Leary, M.; Wege, D. *Tetrahedron Lett.* **1978**, 2811.
(32) Watson, C.; Pagni, R. M.; Dodd, J.; Bloor, J. *J. Am. Chem. Soc.* **1976**, *98*, 2551.
(33) Hafner, K. *Pure Appl. Chem.* **1971**, *28*, 153.
(34) Wang, T.; Matula, R.; Farmer, R. "International Symposium on Combustion, 18th, 1980"; The Combustion Institute: Pittsburgh, PA, 1981; p 1149.
(35) Dias, J. R. *Carbon* **1984**, *24*.

Incorporating Chemical Literature and Information Retrieval into the Chemistry Curriculum at The Hebrew University†

Y. WOLMAN

Department of Organic Chemistry, The Hebrew University of Jerusalem, Jerusalem 91904, Israel

Received September 1, 1983

A chemical literature and information retrieval program has been developed at the Chemistry Teaching Division of The Hebrew University. The program is incorporated in the undergraduate chemistry curriculum for the chemistry major students. The emphasis in the whole program is on the end user, the chemist. The various problems and stages in the program development are discussed. The various parts of the programs are described in detail.

Recently, we, at the Chemistry Teaching Division of The Hebrew University, developed a teaching program in Chemical Literature and Chemical Information Retrieval. This program replaced the traditional descriptive Chemical Literature class that was given to the chemistry major students in their third year.

The whole program is divided into four parts: (a) a compulsory introductory lecture of 4 h followed by a conducted tour of our science library; (b) an elective 10–h class entitled "Chemical Information"; (c) a 4–h lecture dealing with Synthetic Organic Chemistry Literature (compulsory to all students who take the Advanced Organic Chemistry laboratory); (d) using a numeric database for the identification of an unknown (an experiment in the Advanced Organic Chemistry laboratory). In order to understand the development of the program, one should be familiar with the requirements and the curriculum at the Hebrew University, both of which are quite different from the ones in American Universities.

A student at the Hebrew University gets his BA or BSc after 3 years of studies. at present, the academic year at the University is divided into trimesters, each one of 9–10 weeks. In order to obtain a BSc in Chemistry, one has to collect 210 credit points (an hour trimester of a regular laboratory session is 0.5 credit point, an hour trimester of a lecture or an advanced laboratory is 1 credit point). Out of those 210 points, 145 are in chemistry (100 compulsory and 45 elective) and 40 are compulsory in physics, mathematics, and computer programming, while the remaining 25 points are elective ones in physics and/or biology.

DEVELOPMENT OF THE PROGRAM

Until 1977, the Chemical Literature class was given to our students in the 8th trimester under the old classical approach. The students received long lists of primary and secondary sources. No search problems were demonstrated, and very little evaluations and comparisons of the various sources were carried out. Three points guided us in renovating the curriculum: (a) to schedule the class at the earliest possible time, thus enabling the student to make use of the material covered as much as possible; (b) to use a practical approach to the subject by demonstrating in the class various examples and search problems from the various disciplines of chemistry; to teach the student the various available sources and search tools and to compare and evaluate them in selected search history cases, thus emphasizing the subject from the user's viewpoint; (c) to teach the students practical searching by giving them

† Presented at the Chemical Information Section Symposia, 184th National Meeting of the American Chemical Society, Kansas City, MO, Sept 1982.

Table I. Schedule of the Basic Chemistry Lectures for Chemistry-Major Students

1st year			2nd year	
1st trimester	2nd trimester	3rd trimester	4th trimester	5th trimester
general analytical	physical	physical	inorganic organic physical	inorganic organic biochem

search problems to be returned by the next class as well as giving a "take-library" examination at the end of the class.

The earliest possible time for the Chemical Information lecture is in the sixth trimester, as only at the end of the fifth trimester our chemistry major students complete their basic chemistry lectures. As can be seen in Table I, the students receive their basic general and analytical chemistry lectures in the first trimester, their basic physical chemistry lecture in the second, third, and fourth trimesters, their basic inorganic and organic chemistry lectures in the fourth and fifth trimester, and their basic biochemistry lecture in the fifth trimester. Indeed, it was decided to schedule the lecture in the sixth trimester.

After having given the "new" course for the first time, we have found that the upper half of the class enjoyed it in its new form and took active part during and after the lecture hours and their grades were above 85. However, the lower half of the class could not cope with the new approach and got lost somewhere during the middle of the trimester. This was reflected also in their final grades, which were below 65. The next year, we somewhat softened our demands, a change that was hardly reflected in the grades distribution. About two-thirds of the class got 80 and above, while the lower one-third got 55 and below.

A decision had to be taken whether to lower the class level to the level of the lower third, to give the class in two levels, or, maybe, to keep the class level and turn it into an elective course. We decided to change the class into an elective one, and indeed, 30–40% of the second-year students choose the course, those being from the upper half of the students, and their grades at the end of the class are 83 and above (the average grade in the last 3 years has been 89–90).

We were all unhappy with the decision to make the class an elective one. We felt that a chemistry major student should familarize himself with the science library and be able to locate a research article or any other material in the library. A BSc chemistry graduate should know something about *Chemical Abstracts* (CA), *Beilstein*, and *Gmelin* and how to use them. He should know what a scientific journal is and be familiar with the most important ones. Therefore, we have decided to give an introductory 4-h lecture during the first week of the 6th trimester within the framework of the Organic Chemistry laboratory, followed by a conducted tour of the science library. The lecture serves as an introduction to the Chemical Information class for those students that elected the course and is a basic lecture for the ones that for one reason or another did not choose the class.

Faculty members are asking, from time to time, to include in the Chemical Information course specific search topics related to their own classes. We are doing our best to incorporate their wishes either in the form of search examples given in the class or within the take-home exercises. However, we were not able to fulfill the requests of the Advanced Organic Chemistry laboratory teachers to include more information about synthetic organic chemistry literature in the course. Very little of the desired material is covered in the Chemical Information course, and we were unable to enlarge it within the framework of the class. However, a solution was found in the form of a 1-h introductory lecture during the first day of the advanced laboratory.

One of the experiments in the Advanced Organic Chemistry laboratory is the mass spectroscopy experiment. The experiment, which was introduced about 10 years ago, consists of learning the principles of mass spectroscopy (practical as well as theoretical), running a mass spectrum of an unknown compound, and identifying the compound from its spectrum. Having received access to the NIH/EPA Mass Spectra Data Base last year, we decided to incorporate the usage of the database in this experiment.

THE INTRODUCTORY LECTURE

The introductory lecture begins by covering the following topics: communication among the scientific community, the scientific meetings and the scientific literature as a communication media, the scientific journal (the scientific article, the mechanism of scientific publications, trends in scientific publication, the journal of the future), and the science library and its organization.

After a short intermission, we teach the students how to locate needed information and/or sources in and out of the library, starting with how to locate a book or an article once all the bibliographic information is available (using the library catalog, discussing what a reference is, how to write a reference of a book, an article, or a patent) and moving to cases where only a part of the bibliographic data is known (journal indexes and their use) and where little or no bibliographic data are available at all (using CA).

CA is discussed at some length—its historical background (the transformation from a current awareness tool into the most important chemical retrieval source), its coverage (abstracting policy, types of sources covered, etc.), the structure of the weekly issues, issue indexes, volume indexes, and collective indexes, and use of the various indexes for information retrieved. At this stage, we show the student how to locate materials that are not in the library (using union lists and CASSI for journals and union catalogs for books).

The last part of the lecture is devoted to the *German Handbuch* and its concept. *Beilstein* and *Gmelin* are discussed (history, structure, content, arrangments of compounds, indexing systems), and the students are taught how to use these sources. Data obtained through the usage of the *Handbuch* is compared to results obtained by use of CA as a search tool.

The tour of the science library teaches the students more about the organization of the library, they see all the primary and secondary sources that have been discussed in the lecture, and it gives them the opportunity to discuss any unclear points or material related to the subject.

Exercises. The introductory lecture is followed by an exercise. The aim of the exercise is to make sure that the student knows his science library and how to use it. The exercise is given on an individual basis; thus, each student is getting different questions. The student has to locate a book on a specific subject (give bibliographic information about the book and its call number), give information on which volume or volumes of certain journals were published in a specific year, and locate an article by using journal indexes and give a Hebrew abstract of the work. All this requires the student not only to use all that he learned during the introductory lecture (the library catalog, journal indexes, and CA indexes) but also to be able to locate the various journals in the library stacks and to understand a scientific article and/or abstract.

THE CHEMICAL INFORMATION LECTURE

The Chemical Information lecture starts with a 2-h class dealing with creating and maintaining a current awareness program. It covers the following subjects: the importance of keeping updated, creating a current statement (profile) of interest, and the three levels of keeping update (the general or newsy level, the general chemical level, and the specific level). The following topics are covered, and the similar secondary sources are compared: journals to be read or scanned back to back, skimming the current literature (*Current Contents, Chemical Titles*), abstract services (CA—the CA issues as a current awareness tool, nonchemistry abstract sources), standard interest profiles (*CA Selects, Ascatopics, BIOSIS/CAS Selects*), specialized abstract services (*Current Abstracts of Chemistry and Index Chemicus, Current Chemical Reactions, Journal of Synthetic Methods, Chemical Industry Notes*), *Science Citation Index*, patent coverage (*INPADOC Patent Gazette*, Derwent publications, patent coverage by CA), reviews (CA indexes, *Index to Scientific Reviews*), computerized SDI, and arranging and keeping ones own records and files.

Before starting with the searching part, the students received some hints on how to conduct a literature search. Those hints, as well as the ones regarding the arrangement and keeping of records and files, are suggestions and guidelines for the students that will help them locate the desired data they are after in the scientific literature as well as in their own files.

The next 3 h are devoted to the subject of numerical data. The importance of critically evaluated numberical data is discussed, followed by a description of international and national data-collecting organizations and their activity (e.g., CODATA, NSRDS). The locating of the following types of numerical data is discussed: melting and boiling points (using the *CRC Handbook, Heilbron Dictionary of Organic Compounds, Beilstein, Gmelin*, and CA as search tools), crystallography (*CODATA Directory, Landbolt-Bornstein, Molecular Structure and Dimension*), spectroscopy [e.g., ESR, IR, Raman, UV and visible, NMR (^{1}H and ^{13}C), and mass]. Thermophysical and thermochemical data sources are discussed next followed by kinetic data and solubility data sources.

A total of 3.5 h is devoted to chemical reactions, chemical concepts, structural and substructural searching (using CA volume and collective indexes, CA *Index Guide*, CAS *Parent Compounds Handbook*, and *Science Citation Index* as search sources). Search strategies based on CA and SCI are compared and evaluated. The CA index name, IUPAC nomenclature, Wisesser line-formula notation, and CAS Registry Number system are discussed in this part of the lecture.

An hour is devoted to safely and related problems. The primary and secondary sources that cover safety, toxicity, occupational hygiene, and environmental impact and control are covered. Special emphasis is given to general safety practice in the chemical laboratory. Few safety-related searches are conducted within the framework of the class.

In the last half hour, an introductory lecture to computerized searching is given. Computerized searching has been mentioned a few times during the class (while discussing current awareness programs, chemical concepts, sustructural search, etc.). The talk covers the general field of on-line searching with particular emphasis on chemical search.

Exercises. During the Chemical Information class, the students receive three exercises that have to be returned within a week. The first one is given after the first 2 h and deals with building a current awareness program. Each student has to choose two subjects out of a given list of 45, locate a recent review on those two subjects, and describe in detail how he should keep himself updated on these topics, which tools he would use for this purpose, and the advantages of the tools he had chosen over other available ones. The second exercise is given after the numerical data chapter is finished—the students are requested to give two evaluated numerical values (e.g., a reaction constant, a binding constant, ΔH, ΔG, solubility data, spectroscopic data). In case no evaluated data are available, the student is asked to try to evaluate the data he collected (here again each student is getting different questions). The last exercise is given at the end of the class. The students are divided into groups of three to four, and each group receives a manual search problem. Upon return of the results, a computerized search is run by each group, and the computerized and manual results are compared. The search problems that were given in the academic year 1981–1982 are listed in Table II.

Table II. Topics for Comparison of Manual and On-line Search (1981–1982)

(a) LD_{50} of monofluoroacetic acid and its derivatives
(b) ^{1}H and ^{13}C NMR of Oxitocin
(c) analysis of TCDD and other chlorinated dioxins
(d) degradation of oil slicks
(e) determination of aromatic hydrocarbons from engine exhaust in the atmosphere
(f) polywater

Table III. Distribution of Take-Library Examination Topics in the Various Fields of Chemistry (1981–1982)

analytical chemistry	4
applied chemistry	2
biochemistry	8
bioorganic chemistry	9
inorganic chemistry	8
medicinal chemistry	8
nuclear chemistry	2
organic chemistry	8
physical chemistry	9
total:	69

Each of the topics in Table II contains some special aspects such as the comparison with numerical and bibliographic databases (topics a and b), substructure search (topics a and c), comparison between a chemical database—CA—and an interdisciplinary one—*Enviroline, Polution Abstracts, Oceanic Abstracts*, and *Aptic* (topics d and e)—development of a search strategy (topic e), and history of chemical research (topic f).

"Take-Library" Examination. At the end of the class, the students are given their take-library examination. Each of the examinations consists of a scientific article or a news story followed by five to seven questions on the material. In the case of an article, the student also has to give an extended abstract of the work. All answers have to be followed by a full description of the route used to locate the desired information, as well as references to the original scientific work.

In the year 1981–1982, the students were able to choose their take-library examination from 69 different topics in the various disciplines of chemistry. The distribution of the topics within the various disciplines is given in Table III. In the Organic Chemistry discipline, the following topics were given:

Antiaromaticity
Asymmetric Synthesis
Crown Ethers
Helical Molecules
Organometallic Compounds
Polymeric Reagents
Protective Groups
Selective Reductants

Examples to the take-library examinations are given in Figures 1 and 2. Figure 1 is an example of a take-library examination based on a scientific article (*Phys. Rev. Lett.* **1973**, *33*, 1490). The student has to answer the following questions:

(1) Please give an extended abstract of the article.
(2) A Russian group has described at about the same time another route to the formation of element 106. What is their route?

Element 106. A. Ghiorso, J.M. Nitschke, J.R. Alonso, C.T. Alonso, M. Nurmia, G.T. Seaborg, E.K. Hulet and R.W. Lougheed, Phys. Rev. Lett., 33, 1490 (1973).

1) מהם עקרי המאמר ?

2) קבוצת מחקר רוסית תיארה קבלת איזוטופ אחר של היסוד 106. באיזו דרך קבלה אותו ?

3) הדעות חלוקות אם אכן באמת קבלו הרוסים את היסוד 106. קריטריונים לזהוי יסודות חדשים פורסמו ב-1976 ע"י כימאי הגרעין האמריקאים. האינפורמציה שבידי הקבוצה הרוסית איננה מתאימה לקריטריונים אלו . מהם הקריטריונים בהם מדובר ?

4) הכימאים התיאורטים מצפים למה שנקרא island of stability עבור היסודות הסופר כבדים בסביבות יסוד 114. מי פתח תיאוריה זו ?

5) נעשו מספר נסיונות למצא יסודות סופרכבדים בטבע. תאר נסיונות אלו.

6) היתה טענה (שהוכחה לאחר מכן כלא נכונה) כי אכן סונטז היסוד 114. תאר את החלק הנסיוני של עבודה זו.

Figure 1. Example of a take-library examination.

1) מי בודד לראשונה את ה-interferon ?

2) מה ידוע לך על פעילותו האנטי וירלית והאנטי סרטנית של ה-interferon ?

3) כיצד מפיקים היום interferon ומי הם בעלי הפטנט על הפקתו ?

4) מה זה clones ?

5) מהם תחומי התעניינותו של Weissmann ?

6) מה ידוע לך על חברת Biogen ?

7) מהו הפירסום המדעי של החומר שהתפרסם לראשונה ב-Time ב-28.1.80 ?

Science

Time – 28/1/80

Genetic Coup

E. coli makes interferon

Almost from the time of its accidental discovery by scientists in England in 1957, interferon has been the stuff of researchers' dreams. A complex bodily protein, it possesses both antiviral and antitumor properties, which means it could become an important new weapon against a wide range of diseases, from the common cold to certain cancers. But it takes 65,000 pints of blood to get just 100 mg (.0035 oz.) of the protein, so testing of the possible miracle drug has been severely limited. Now, as a result of another application of gene-splicing, or recombinant DNA, techniques, all that may change. In Boston last week molecular biologists announced that they had induced tiny bacterial "factories" to copy human interferon.

Ordinarily, interferon is produced by virtually all cells in the human body, which get their instructions for making it from a specific gene in their DNA; these are passed on to the cells' protein-manufacturing sites by a genetic molecule known as messenger RNA. But for Hungarian-born Charles Weissmann of the University of Zurich, and his Swiss, Finnish and Japanese colleagues, the natural process was only a starting point. After extracting messenger RNAs from human white blood cells, which were producing interferon, they used these molecules to generate sections of DNA that they hoped would include the required gene. They then spliced these fragments into the genes of a laboratory strain of *E. coli* bacteria. The bacteria began to make copies, or clones, of their altered selves.

Eventually, the team produced some 20,000 clones. But which ones carried the crucial interferon DNA? Analyzing them in successively smaller groups—first 500 at a time, then 64, then eight—the scientists isolated their genetic needle in the haystack: the bacteria that carried the DNA for interferon. After they found one such bug, they could easily identify others and extract the DNA fragments. The team spliced them into different places in *E. coli*, and, presto, the bacteria began cranking out a close facsimile of the human protein.

The amount of interferon made by these bacterial minifactories was extremely small and impure. But the researchers, who did their work on behalf of Biogen S.A., a Swiss-based firm set up to exploit recombinant DNA technology, hope to produce larger quantities of purer material. The ultimate goal: to bring down the cost of an injection of interferon from today's price of $75 a shot to as little as $1. ■

Figure 2. Example of a take-library examination (Reprint with permission from *Time.* Copyright 1980 Time-Life Inc.).

(3) It is questionable whether the Russian group really obtained element 106. New criteria for the identification of new elements were reported by the American nuclear chemists in 1976. The date reported

by the Russian group do not fulfill those criteria. What are those criteria?

(4) The theoretical chemists are expecting "island of stability" for super-heavy elemnts around element 114. Explain why.

(5) A few experiments have been carried out to locate super-heavy elements in nature. Please describe them.

(6) There was a claim (which turned out later to be wrong) for the synthesis of element 114. Describe the experimental part of the above work.

Figure 2 is an example of a news story that was published in the Science section of *Time* magazine on January 28, 1980. The student has to answer the following questions:

(1) Who was the first one to isolate interferon?

(2) What do you know about the anti-viral and anti-carcinogenicity activity of interferon?

(3) How is interferon isolated today from natural sources? Are those routes protected by patent(s), and if so, who holds the patent(s)?

(4) What are clones?

(5) What the research interests of Dr. Weissmann?

(6) What do you know about the Biogen Co.?

(7) What is the scientific work on which the *Time* article is based?

SYNTHETIC ORGANIC CHEMISTRY LECTURE

The Synthetic Organic Chemistry Literature lecture covers the following topics: locating suppliers of fine chemicals, locating procedures for synthesis of chemical compounds (*Beilstein*, CA, *Organic Syntheses*), designing a synthesis (the use of organic class preparations sources such as *Methodicum Chemicum*, *Houben-Weyl*, *Compendium of Organic Synthetic Methods*, *Formation of C-C Bonds*, *Theilheimers Synthetic Methods of Organic Chemistry*, *Journal of Synthetic Methods*, *Current Chemical Reactions*), information concerning reagents, solvents, and/or techniques (*Reagents for Organic Synthesis*, *Techniques of Chemistry*), and getting an overview (*Organic Reactions*, *The Chemistry of the Functional Group*, the review literature).

USAGE OF NUMERICAL DATABASES

The last part of the program is the usage of a numerical database in the advanced laboratory. The students received a short introductory lecture about data banks in general and numeric data banks in particular. The structure and organization of the NIH/EPA MSSS (Mass Spectral Search System) is discussed. The student learns about some of the database options and their usage [e.g., searching for peaks (PEAK), searching for peaks and molecular weight (PMW), entering a spectrum, searching with a complete spectra.] After running the mass spectra of the unknown, the student elucidates the structure of the compound and then turns to the terminal in order to run a computerized structure elucidation, comparing the results obtained from the two routes.

CONCLUSIONS

Our feelings are that a student that participates in all four parts of the program receives a good basis for the modern approach in chemical information. The ones that take only the compulsory 4-h lecture receive the basic information about the chemical literature. Those students learn some of the aspects of chemical information (the use of numerical databases and synthetic organic literature) if they elect the Advanced Organic Chemistry laboratory.

Quantification, Retrieval, and Automatic Identification of Numeric Data in Organic Chemistry Journals

JOSEPH J. POLLOCK

Chemical Abstracts Service, Columbus, Ohio 43210

Received December 20, 1983

The full text of American Chemical Society (ACS) primary journals may now be searched on-line. An important benefit of this capability is that it makes available a substantial amount of highly current numeric data, which can be used to help identify unknown compounds or retrieve data on known ones. Such searches are readily performed with high recall and precision. The amount of certain types of numeric data is estimated, and an algorithm for identifying them in the journal text is discussed in detail.

INTRODUCTION

Computer searching of the content of chemistry journals has traditionally proceeded via index entries and focused on concepts rather than data. However, computer technology has now advanced to the point where it is possible to search primary journals directly and for data. A particular aspect of this topic is explored here; the quantity, use, and automatic identification of numeric data in the experimental sections of three American Chemical Society (ACS) journals concerned with synthetic organic chemistry. All ACS journals can now be searched on-line in a publicly available file.

The first half of this paper estimates the amount of characterization data (defined below) in three ACS journals and demonstrates how one can use it to identify an unknown compound or retrieve data on a known one. The second half presents an algorithm for identifying a "package" of numeric data in primary journal text comprising a CAS Registry Number (CAS REG) for a substance, its name and/or symbol, and its characterization data together with their measurement conditions.

(1) QUANTIFICATION OF CHARACTERIZATION DATA

When ACS journals were examined for numeric data, it rapidly became clear that the kind most likely to repay further study was "characterization data", the type that chemists

0095-2338/84/1624-0139$01.50/0