A One-Quarter–One-Credit Course in Chemical Literature†

JACK D. GRAYBEAL

Department of Chemistry, Virginia Polytechnic Institute & State University, Blacksburg, Virginia 24061

Received November 23, 1983

A one-quarter–one-credit course in chemical literature has been designed to fit into a crowded chemistry curriculum and still give the student a meaningful introduction to the subject. Following a discussion of the objectives and general nature of the course, a brief synopsis of each lecture and attendent exercises is presented. Although all desirable topics cannot be covered, experience has shown that the course does achieve the primary objective of providing sufficient background to enable the student to perform independent literature searching.

INTRODUCTION

At Virginia Polytechnic Institute & State University, a one-quarter course in chemical literature is used to provide the students with a meaningful background in the literature within the available time frame of 10 h of lectures. The important arguments related to the need for the establishment of the course are first reviewed to indicate the environment within which the course was developed. Next, the general nature and philosophy of the course is presented. The detailed contents of each lecture and the attedent exercises are given. Finally, conclusions regarding the success of the course are stated.

BACKGROUND

The development of the course took into account the fact that it was to be taught to students who were at the beginning of their junior year and who would not have the chemical background of either seniors or graduate students. When the faculty first began discussions leading to the initiation of this course, there was a division of opinions as to how best to teach the use of the literature: whether to use a formal course structure or to teach it within the structure of undergraduate research. The use of undergraduate research as a means of teaching the chemical literature generally restricts the breadth of the material that is covered, whereas the use of a formal course necessitates that some other area of the chemist's training be reduced in order to satisfy total course requirements of the College of Arts and Sciences.

The present course is a compromise between the amount of time that is desirable to cover the subject and the time that the faculty was willing to devote to it in the curriculum. The course is taken by students during the first quarter of the junior year and is offered at the earliest possible time so the students can get the maximum use of the training prior to graduation. It is followed in the senior year with a seminar that requires the students to make further use of the literature.

GENERAL NATURE OF THE COURSE

There are four primary reasons for requiring this course. First, the student needs to learn techniques to increase his conceptual knowledge, thus enabling him to gain a greater understanding of existing concepts and to develop new ones. Second, the course provides the student with the background to retrieve data, as well as methods, necessary to enhance his ongoing work. A third and very important aspect is training the student to prevent duplicating work that has already been undertaken. The development of proper techniques regarding literature use can shorten the time necessary for a person to retrieve material from the literature. Also, there is no need to reinvent the wheel. Nothing is more damaging to one's professional stature than to submit a manuscript describing a new synthesis or technique only to have a referee point out that it was published in *Berichte* in 1896. Finally, the student needs to be aware of the problems associated with keeping current in the many areas of chemistry, a very difficult task if one is to remain at the forefront of the field.

The course is limited in time, therefore necessarily limited in scope. There is little time to teach chemistry, so care must be taken to ensure that the chemical knowledge required for a particular exercise does not exceed the student's background. The instructor must point the way; only highlights and principal concepts can be covered in lectures, and the student is referred to texts and supplemental material for more detail. The exercises must be carefully chosen and checked to be sure that they are suitable. What may appear to be a good problem on first thought can often turn out to be too complex or too difficult for the student at his level of schooling. The exercises must be diversified to cover the maximum number of concepts. When the class size becomes large, parallel exercises must be used in order to prevent overloading of the use of single documents. It is emphasized to the student that the technique for learning is different for a chemical literature course than it is for a conventional chemistry course in which the student can sit down at the desk with his text and problem set, work through the problems, and gain a reasonable understanding of the material. In the study of the chemical literature, there is absolutely no way to learn other than to go to the library and search for and use the material.

The course consists of four primary facets. (1) There are 10 lectures. In the course of the lectures, references to important sources of information regarding the use of the literature are discussed. (2) There are nine exercise assignments, each of which entail, for the average student, 2–3 h of work. (3) The students are asked to do a short retrospective search. In selecting the topic for this search, it is necessary to keep the subject narrow and restrict the time period over which the search is conducted. (4) The course is terminated with a final examination that contains questions regarding major sources of information and the use of these sources.

The general content of the course consists of an overview of the nature and the types of chemical literature. Initially, the basic types of chemical literature are discussed and ca-

† Presented before the Division of Chemical Information and Chemical Education, Symposium on Chemical Literature and Information Retrieval in the Chemistry Curriculum, 184th National Meeting of the American Chemical Society, Kansas City, MO, Sept 15, 1982.

0095-2338/84/1624-0193$01.50/0

tegorized, and then examples are shown to the students. Following this, some of the more important sources, as well as thost that are more difficult to use, are selected and discussed in some detail. The general problems of using the literature and conducting a search are reviewed. It is emphasized to the student that (1) he should build a search from a general base obtained by consulting texts, monographs, and review articles, (2) there is a need to organize the search before starting, and (3) the selection of proper terms for setting up a search is essential.

DISCUSSION OF LECTURES

The first lecture is designed to give the students a general introduction to the chemical literature, disciss the organization of the on-campus library where they will be working, provide an introduction to the methods used for obtaining information about a specific topic, and provide instructions on obtaining specific materials. The exercise associated with the first lecture requires the student to go to the library, use the library catalog, locate materials on the shelf, and take a cursory look at the material in order to decide on its type.

The second lecture discusses the use of indexes and the role that these play in current awareness through publications such as *Chemical Titles* and provides an introduction to abstracts. At this time, a general introduction to *Chemical Abstracts* with particular emphasis on the use of individual issues is presented. The assignment associated with this lecture is designed to compare the coverage of *Chemical Titles* and *Chemical Abstracts* and introduce students to using current issues of these publications.

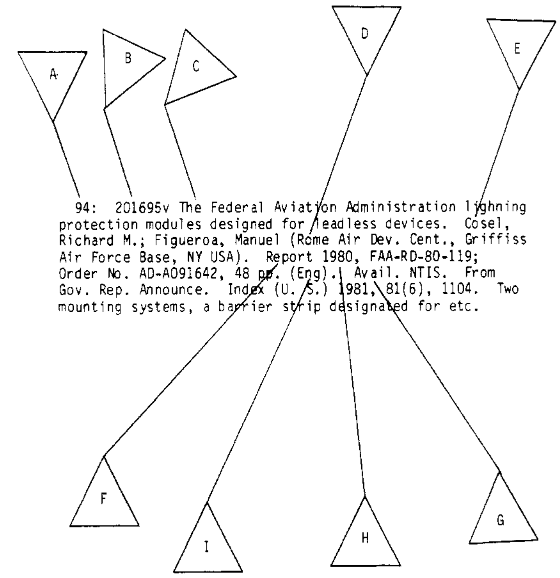
The third lecture is a discussion of the use of annual and collective indexes to *Chemical Abstracts*; the associated exercise requires the student to locate very specific items within a limited time frame.

Lecture four is designed to acquaint the student with the supplemental aids for use of *Chemical Abstracts*.[1,2] It concentrates on the nature and utility of these aids rather than on the detailed use of any particular one. This is necessary due to the time limitations. The associated assignment requires the student to locate more obscure entries that necessitate the use of the aids. This is an area in which great care must be exercised in selecting assignments so as not to get the student lost.

Lecture five has its primary emphasis on *Current Abstracts of Chemistry* and presents a brief set of comments on other available abstracts. The exercise assigned here is designed to have the student compare entries relating to the same journal articles as found in *Current Abstracts of Chemistry* and *Chemical Abstracts* and to look at the advantages and disadvantages of each publication.

Lecture six looks at the problem of tracing the chemical literature through the use of author entries and focuses on the use of the *Science Citation Index*. Additionally, there is a discussion of general information sources. The accompanying assignment gives the student the task of tracing the work of a particular author through the literature and the problem of finding answers to general questions in encyclopedias, dictionaries, and handbooks. This exercise is designed to require the student to use a variety of general information sources rather than to obtain all answers from a single source.

In lecture seven, emphasis is placed on a careful search for review articles and the importance of these articles in providing a base for extended research. It is pointed out that the review articles are very good for establishing search terms, locating key references, and preventing duplication of effort. Important sources of chemical data are reviewed, and the accompanying assignment gives the student a topic on which to locate reviews and find specific data.



A. CA volume number
B. CA abstract number
C. Title
D. Place where work was done
E. Authors
F. Source of article
G. Availability of article if not a conventional publication
H. Year of publication of article
I. Language of article

Figure 1. Typical abstract.

Lecture eight is devoted to a review of the minor compendia in the various areas of chemistry and a somewhat more detailed discussion of the use of *Gmelin*. The associated assignment is aimed at getting the student to use *Gmelin*, in spite of the fact that he may lack any knowledge of German.

Lecture nine covers high points in the use of *Beilstein*, and an assignment provides the students with limited use of this compendia.

The final lecture discussed the available computerized data bases and methods for accessing them. Emphasis is placed on the need for the student to know the field and to be able to search manually prior to trying a computerized search. Illustrations of computer searches are presented to the class, and the economics of computer searching is discussed. The need for careful selection of terms to keep output within reaon is also pointed out. There is a general discussion of the problems associated with trying to keep abreast of the output in the field and with the methods of so doing. The use of Chemical Abstracts Service and interlibrary loans as a source of material is discussed.

Several books on chemical literature were available when the course was initiated,[3-6] and a new edition of Mellon[7] has recently been published. The text by Maizell[3] was used for the first two quarters in which the course was taught.

Throughout the entire course, extensive use is made of audiovisual aids. These aids are almost exclusively constructed in house by using anotated excerpts from the literature and combined in a manner so as to illustrate the maximum number of important features of the source being discussed. Two examples illustrating the nature of an abstract from *Chemical Abstracts* and the relationship between an abstract and an author index are given in Figures 1 and 2. Another valuable source of material has been found to be trade publications such as those published by Chemical Abstracts Service,[8,9] Beilstein,[10] and Gmelin.[11]

CONCLUSIONS

Now that this course has been taught twice, there are a number of important conclusions that we can reach. The available time is too restricted; one can only do a minimal, not

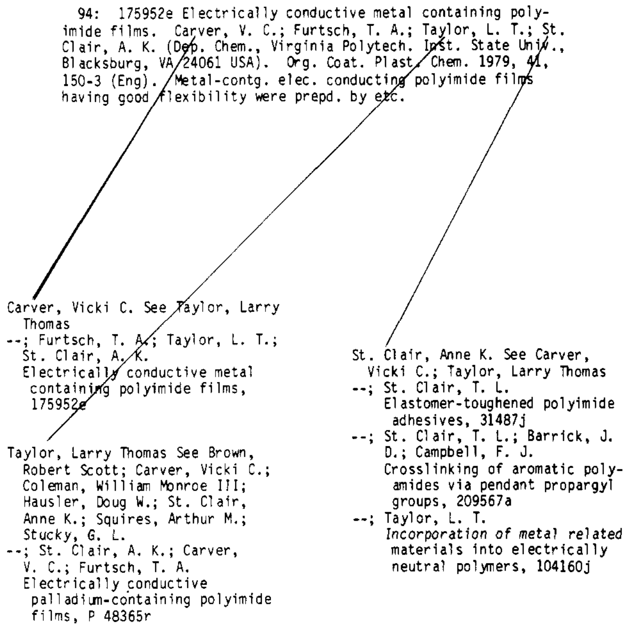


Figure 2. Relationship between author index entries and abstract.

adequate, job without overloading the students with outside work. Even with the restricted nature of the coverage, the work time required of the student is still relatively high for a one-quarter course. Several important areas do not receive the in-depth treatment that they really need. Such areas include patents, government documents, *Chemical Abstracts* supplemental sources, strategies for searching, and the use of *Beilstein*. It would be possible to change the order in which the material is presented, discussing some of the more general sources of information first, which would be somewhat more logical. However, this would limit the use of a retrospective search as an assignment in the course. The overall response of the students is that while they found the course to be very time-consuming, they, indeed, consider it to have been very beneficial to their training, and despite the restricted coverages, it served the purpose of providing a base from which they can operate in the future.

REFERENCES AND NOTES

(1) "Chemical Abstracts Service Source Index", CASSI; Chemical Abstracts Service: Columbus, OH, 1979.
(2) "Chemical Abstracts Index Guide"; Chemical Abstracts Service: Columbus, OH, 1982.
(3) Maizell, R. E. "How to Find Chemical Information"; Wiley: New York, 1979.
(4) Bottle, R. T. "Use of Chemical Literature"; Butterworths: London, 1979
(5) Herner, S. "A Brief Guide to Sources of Scientific and Technical Information", 2nd ed.; Information Resources Press: Arlington, VA, 1980.
(6) Woodburn, H. M. "Using the Chemical Literature"; Marcel Dekker: New York, 1974.
(7) Mellon, M. G. "Chemical Publications", 5th ed.; McGraw-Hill: New York, 1982.
(8) "Information Tools"; Chemical Abstracts Service: Columbus, OH, 1978.
(9) "Searching CA"; Chemical Abstracts Service: Columbus, OH, 1981.
(10) "How to Use Beilstein"; Springer-Verlag: New York, 1981.
(11) "This is Gmelin"; Springer-Verlag: New York, 1981.

Condensed Structure Identification and Ring Perception[1]

JAMES B. HENDRICKSON,* DAVID L. GRIER, and A. GLENN TOCZKO
Edison Chemical Laboratory, Brandeis University, Waltham, Massachusetts 02254

Received February 10, 1984

For acyclic graphs, the full connectivity is uniquely represented by a T-list of the row sums from the maximal adjacency matrix, requiring only $2n$ bits in length. Polycyclic graphs are represented by the T-list of the maximal spanning tree as well as an R-list of the row sums of ring-closure bonds on that tree, also requiring only $2n$ bits for storage. The combined T/R-list ($4n$ bits) provides for unique reconstruction of the graph in most cases, and in all cases when an offset number (ON) is appended to identify the rank order in the few instances for which duplication occurs. A procedure for reconstructing the matrix, and hence the graph itself, from the T/R-list (and ON) is presented. A rapid protocol for perception of the smallest set of smallest rings (SSSR) in the graph also derives from the maximal matrix. All the procedures are contained in a program (TRGEN) written for minicomputer.

We recently showed that a chemical structure, or any graph, can be uniquely numbered by creating its maximal adjacency matrix.[2] This maximal matrix is created by assigning numbers to the atoms, or graph points, in such a way that each row in the matrix, considered as a binary number, must be the maximum possible number. The full binary number representing the molecular skeleton is obtained by stringing out each successive matrix row of 1/0 entries to the right of the diagonal into a single list, to be seen as a binary identification number for that skeleton. This number is unique because it is the maximum possible binary number among all the possible $n!$ matrices representing different numberings of the skeleton. When the skeleton, or graph, has isomorphic atoms or points, the same maximal matrix represents each equivalent numbering, and it was shown that such isomorphic points can be easily identified via equivalency classes.[2]

The nature of the maximal adjacency matrix has some further properties, explored here, which allow an unusually condensed form of representation of the matrix for compact storage and also a very simple basis for determining the smallest set of smallest rings (SSSR)[3] in the molecule. In general, the adjacency matrix is a symmetrical $n \times n$ matrix, of 1 entries indicating bonds or connections between atoms and of 0 entries for no connection. Since the maximal matrix is created row by row to make each row a maximum binary number, the effect is to press the 1 entries as far up and to the left as possible (above the diagonal) and so to leave an extensive region of 0 entries to the upper right in the matrix. It is this effect that allows a much shorter representation of the matrix to be written and also allows the smallest rings to be identified readily and rapidly.

Acyclic Molecules: The T-List. Let us take first the case of the acyclic molecules, or tree graphs. The procedure for creating the maximal matrix is to assign the numbered rows in the (empty) matrix successively to atoms, in such a way as to make each row a maximum. This requires that each row

