the search to get what is wanted, as the patents retrieved do not suggest useful additional keywords for searching on.

An average search takes five minutes to superimpose and read off the coincidence cards, and a further 15 minutes to pull out the papers themselves. The system is used two or three times per week. As in the previous system the number of answers may be increased or decreased by deleting or adding keyword cards. Because the file is small and contains only material of interest to the establishment in which it operates, retrospective patent searches are quicker than with *Chemical Abstracts*. Moreover, within the establishment's limited field of interest, the system contains more material, and perhaps of more importance all the patents retrieved are available close at hand. The system is easier to use and gives greater precision in searches for British specifications than the British patent office classification scheme. This is because the keywords are tailor made for the type of searches which are carried out. However, if an exhaustive search is needed, particularly in a field of interest new to the establishment, the patent office scheme and *Chemical Abstracts* should also be used, as the internal system may not contain all relevant patents in this new field. Derwent patent publications and the official indexes and abridgements to foreign patents are useful when searching for overseas patents, but the latter publications are often difficult and tedious to use, and it is hoped that the inclusion of patents from other countries in the present system will make these searches easier.

The effectiveness of the presently described patent retrieval system or, for that matter, of any retrieval system depends ultimately on the contents of the file. That which is not there cannot be retrieved. In the particular case under consideration, it depends on the efficiency of the establishment in selecting patents of current interest, and of foreseeing changes in interest.

ACKNOWLEDGMENT

LITERATURE CITED

(1) Hunter, P. S., "Information Retrieval in the Technology Based Industries," *Chem. Ind. (London)*, 1971, pp. 667–70.

(2) Hunter, P. S., "The Foreign Language Barrier—Stumbling Block or Stepping Stone?" *Inform. Sci.* 4, 65–70 (1970).

(3) McGarvey, A. R., "Uniterm Index to U. S. Chemical Patents —User Evaluation," *J. Chem. Doc.* 8, 23–5 (1968).

(4) Silk, J. A., "A Notation-Based Fragment Code for Chemical Patents," *Ibid.*, 8, 161–5 (1968).

(5) McDonnell, P. M., "Technical Information Management in the U. S. Patent Office," *Ibid.*, 9, 220–3 (1969).

(6) Rasmussen, L. E., and Van Oot, J. G., "Operation in DuPont's Central Patent Index," *Ibid.*, 9, 201–6 (1969).

Computer Aided Bibliographies for Personal or Group Use

R. J. CEDERGREN

Département de biochimie, Université de Montréal, Case Postale 6128, Montréal, Québec

A computer-based personal bibliographic system is described. Input via punched card containing the main author, keywords, reference, a number corresponding to the reprint list, and a commentary yields outputs consisting of alphabetized listings of authors and keywords.

Recently computer technology has been developed to aid the scientist in handling the mass of printed information. For example, the Canadian National Library[1] offers a Selective Dissemination of Information (SDI) based on tapes obtained from Chemical Abstracts Service and the Institute for Scientific Information. The computer search may be done on the basis of keywords, author, or references cited by the author. The availability of these alerting systems incites the development of equally rapid and work-saving personal or group reference systems.

More recently a number of computer-based systems have been developed. One type involves a computer search of stored material along the same lines as the commercial computer searches outlined above.[2-5] Unfortunately, this system requires that a computer be available at the time one is looking for a reference. When only large computer installations are available, the time needed to make the search and receive the output must be considered in the over-all efficiency. These types of search are, however, very useful for preparing a bibliography for a new member of a research group, for example. In larger groups where a small computer is available on demand, tremendous amounts of literature can be facily handled to great advantage by this type of search.

A second type of computer treatment involves the printing of lists of references generally arranged alphabetically.[6] This method offers the following advantages: no need for a computer after the list is made, the ease of reference recovery by scanning a printed list, and no need for complex peripheral equipment. There are, however, certain disadvantages—namely, the computer and printing time in preparing the lists, which necessarily limits the number of entries to perhaps a few thousand references (thereby limiting its usefulness to smaller research groups).

This paper describes a system based on the second type of treatment. This system has been used for the past year and a half by our research group comprising 5 to 7 people. Because of our interests in a rather prolific field, t-RNA structure, sequence and evolution, the development of the described system was judged a necessity.

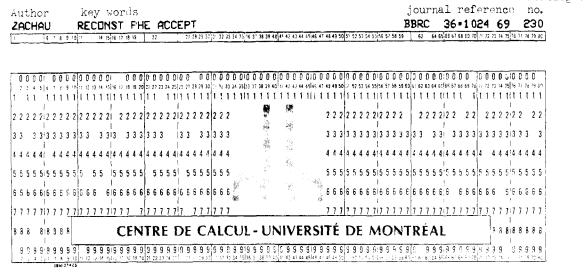# COMPUTER AIDED BIBLIOGRAPHIES FOR PERSONAL OR GROUP USE

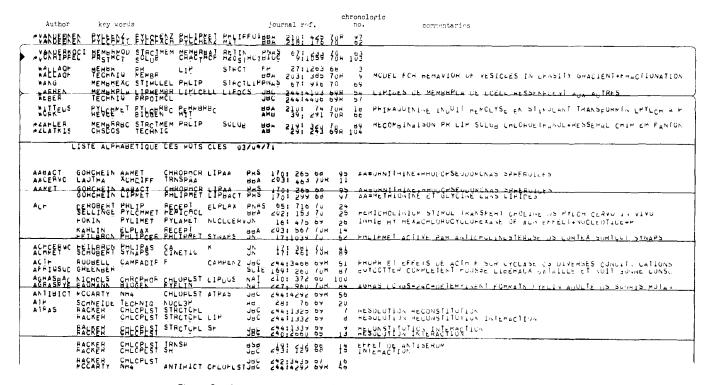Figure 1. A typical reference card to be used with program CHERCHE.

Figure 2. Alphabetized listing of authors outputed by program CHERCHE.

THE OVER-ALL VIEW

We currently use the Canadian National Library reference alert system which sends each week 10 to 40 references. Around 50% of these references can be eliminated by a casual verification. The remaining references after being divided up by the research group are reviewed for pertinence. Here again certain references are discarded, and the reviewer is then responsible for securing a reprint of the remaining articles. The reprints are collected approximately every month and are coded into the personal search program, CHERCHE, described below. Coding is followed by card punching and a list updating. Thus, an up-to-date reference list is available with an average lagtime

(date of publication–date of listing) of approximately two or three weeks.

It is important to emphasize the fact that after this lagtime, all people in the group have been alerted and have access to all pertinent articles on the subject of interest.

PROGRAM CHERCHE

Program CHERCHE is written in Fortran IV for a CDC 6400 computer. A listing of the program with details for its use is available on request. References are entered into tab cards in the following fashion. The first 10 spaces are reserved for the primary author. Spaces 11 through 59 are available for up to six keywords or subject classifi-

cations of 8 or fewer letters in a free format style. The free format is permitted by the program's word construction routine. A new word is recognized by the space between the words. By using words, coding dictionaries of subject classifications are eliminated.[3] Spaces 60 through 75 are used for the journal reference and 76 through 80 for the reference number chronologically determined corresponding to the place in the communal reference file. A second input card can be used for comments or the title of the article. This card is not analyzed by the computer, but is printed automatically with the information on the first card. Keypunching of a data card takes about 30 seconds.

The output consists of an alphabetized list of authors complete with keywords and comments followed by an alphabetized list of all keywords complemented by the rest of the stored information on the reference where the key is found. Therefore, an article with four keywords appears five times on the complete reference list—one time for each of the keywords and once for the author. References which have not been treated previously are marked with an arrow to distinguish them from old references. This fact as well as the number of copies of the listing are indicated on the first input card. The running time of the program, currently with 600 references (1200 input cards), is on the order of 2 minutes for two copies of output.

The output of CHERCHE has the advantage over KWIC and KWØC outputs in that only the information required by the group which may not be included in the title is tabulated. KWIC with enriched titles approaches the utility of the system described here; however, it is a more complicated treatment and gives more unusable output coming from general terms used in titles.

PRACTICAL USE AND FUTURE OF CHERCHE

Currently we use four keywords per reference for economy of input and output. While some may think more keywords would be useful it should be realized that this reference is automatically cross-referenced four times, which is an improvement over conventional systems. We have found that part of the restrictiveness of four code words can be eliminated by using complex words. For example, 5S-RNA and 16S-RNA can be inputed under the form RNA5S and RNA16S which when printed will be listed at the end of the articles having the key word RNA. Two advantages are gained—the complex word gives more information than one simple word, and all articles on all types of RNA are classified together.

Another problem arises in using a keyword which is too general. MEMBRANE, for example, is likely to give too many entries making it difficult to find a particular article using this code word. There are two ways to reduce the number of articles under a particular keyword: by using complex words as suggested above (an example might be MEMBMIT which refers to mitochondrial membrane) and by not using the particular term at all. In our case, 90% of our articles deal with t-RNA. The use of t-RNA as a key word, therefore, is not necessary.

To summarize we have developed an integrated literature system convenient for small groups (1 to 10 people) which has the following features:

A quick reference alert system

An internal alert and classification system based on program CHERCHE

A reference system which permits and encourages a division of labor

A chronologically arranged reprint and photocopy file which is available in the laboratory for quick consultation

We hope that by further modifications, such as magnetic tape storing and updating procedures, computor time can be reduced to a minimum. Microfilm storage of reprints, could also be advantageous by reducing the recovery time for articles stored in the reprint file.

ACKNOWLEDGMENT

LITERATURE CITED

(1) Brown, J. E., "The CAN/SDI Project," *Special Libraries* **60**, 501-9 (1969).

(2) Dewi, P. J., and Braet, W. W., "Computer Documentation System for Small and Medium-Sized Information Collections," *J. Chem. Doc.* **10**, 95-7 (1970).

(3) Gillis, C. N., "Biomedical Information Retrieval: A Computer-Based System for Individual Use," *J. Chem. Doc.* **7**, 98-100 (1967).

(4) Jameson, D. L., "Information Retrieval for the Working Scientist: A Simple Algorithm," *BioScience* **19**, 232-3 (1969).

(5) Arnett, E. M., "Computer-Based Chemical Information Services," *Science* **170**, 1370-6 (1970).

(6) Bridges, K. W., "Automatic Indexing of Personal Bibliographies," *BioScience* **20**, 94-7 (1970).