Conversion of BIOSIS Previews Database Files To Reference Text Files and Reprint Request Forms

John A. Byers

ABSTRACT   **BIOSIS.EXE** is a compiled BASIC language program for use on IBM-compatible personal computers. The program converts a DOS text file of current BIOSIS Information Transfer System (B-I-T-S) references into a DOS file of standard scientific citations. The latter file can be imported into a word processor or scientific reference manager. The program also makes a random access database of references consisting of a title, author, journal, author address, and key words. This database can be searched for specific references and these selected for conversion into a DOS text file of citations as above. The random access database is also used to construct reprint request cards from information on the author, title, journal, and address as well as the user's name and address. The reprint request cards can be printed on regular paper or on "Request-a-Print" (Institute for Scientific Information) cards. The program also can process DOS files obtained from key word bibliographic searches of BIOSIS Previews (which differ in file format) to obtain scientific citations for word processing or reference managers.

TODAY'S SCIENTIFIC LITERATURE is voluminous, fast changing, interdisciplinary, and accrues at an ever-expanding rate. Researchers and teachers of entomology thus may have difficulty keeping up-to-date with the pertinent literature. During one's career, new avenues of study are explored and thus one must compile databases of scientific references in addition to any that have been previously maintained. Fortunately, there is the personal computer that can be used to store, manipulate, and retrieve references from a database. Commercial database services such as BIOSIS Information Transfer System (B-I-T-S) can provide the latest references on computer disk for viewing on personal computers. Another computer database, BIOSIS Previews, can often be used at a university library for retrospective searches on a particular topic. The references located by such a search of the many bioscience journals and publications, from the present back to 1970, can be downloaded to disk for a fee.

I have developed a compiled BASIC program, **BIOSIS.EXE,** that runs on IBM-compatible computers from the DOS command line. The program can take DOS text files of either current monthly B-I-T-S listings or of retrospective searches of BIOSIS Previews and make a database. This database can then be searched for key words and the located references marked for export to a word processor in a format resembling a standard scientific citation. The marked references can also be formatted and printed as reprint request cards. Other commercial databases can be imported such as MEDLINE (National Library of Medicine) and CA Search (American Chemical Society), but exported citations may not be formatted entirely accurately.

Materials and Methods

The program **BIOSIS.BAS** was developed with the Quick-BASIC programming language (version **4.5,** Microsoft). The source code was compiled to **BIOSIS.EXE,** a binary machine-coded executable file, by the Microsoft compiler version **4.5** and overlay linker version 3.69. IBM-compatible computers can execute the program in monochrome or color using any type of monitor and DOS 2.1 or higher.

An IBM-DOS text file from BIOSIS Previews (retrospective search) or B-I-T-S (monthly update) is required for processing by the program. Files from retrospective searches can be downloaded from a number of commercial computer databases such as Data-Star, ESA-IRS, DIALOG, and STN. References in a BIOSIS Previews text file may contain the subject identification codes (field tags) of AN (accession number), AU (author), TI (title), SO (source), IN (author affiliation), KW (key words), BC (biosystematic codes) and ST (supertaxa). Other codes may include (CD, LG, YR, CC, and AB). Monthly updates of recently published articles are formatted slightly differently by BIOSIS such that the AN code is replaced with an ID (identification) field code and the CS (corporate source, author address) may substitute for the IN code above. The use by the program of different field codes in BIOSIS files as well as other databases is explained subsequently.

Results and Discussion

The major facilities and operation of **BIOSIS.EXE** are diagrammed in Fig. 1. The arrows between boxes indicate relationships between program parts but do not necessarily denote program flow. **BIOSIS.EXE** is used by typing *biosis* at the DOS command line. The program loads and checks to see if a name of a database is in the file **BIOSIS.DBD.** If no name is present then the default name (**BIOSIS.BIO**) is placed in the file. The menu has seven options to choose from (Fig. 1) besides a help section, shelling to DOS, or exiting from the program.

The first option is to add BIOSIS files to the database. This is done by entering the name of the BIOSIS Previews file (**e.g.,** named BIOS-1 and represented in Fig. 2A). The user may look at the BIOS-1 file to see if the formatting is such that all text is indented three spaces except for the codes. If the file has been corrupted, a reformatting facility is selected. Otherwise, the user continues and is presented with a list of files (ending in .CD#) that contain data for the various field codes. The user selects the appropriate file name, if present, and the codes are loaded and presented on a menu for editing. The codes can be used without editing or new codes added for the respective authors, title,
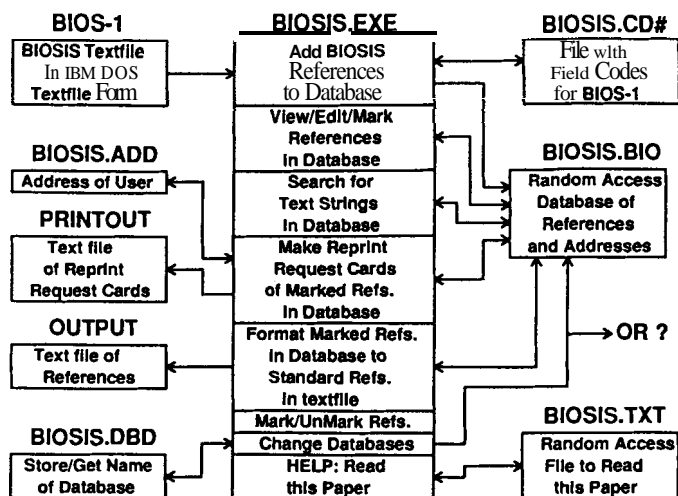
Fig. 1. Relationships of program parts of BIOSIS.EXE, a BIOSIS Previews file conversion utility and reference management and reprint request card system for the IBM-compatible personal computer.

journal, and so forth. These codes make it possible to accommodate any of several types of BIOSIS formats as well as other database styles such as CA Search and MEDLINE. After editing the field codes, the BIOS-1 file is used as the source of references for placement in the BIOSIS.BIO file. This latter file is a random access database in which references from additional BIOSIS Previews files can be added by running option one again. Alternatively, another name can be chosen in order to create additional databases or the BIOSIS.BIO file can be erased to begin a new

database with this name. Each record of the database has the reference reorganized so that the title is placed at the top of a 12-line by 80-character area, followed by an author symbol and the authors (Fig. 2B). The journal or publication is third followed by the author address and key words if present (Fig. 2B).

The second option "View/Edit/Mark References in Database" allows one to scroll through the database and view the references. Any references that are desired to export as a citation in the file OUTPUT or as a reprint request card in the file PRINTOUT are marked by pressing the *Insert* key. References are unmarked with the *Delete* key. A global unmarking of all references can also be done. Facilities for searching for a text string and viewing the located reference are provided while viewing the database.

The third option, "Search for Text Strings in Database," allows complex logic searches of the database fields for up to 2 sets of 20 text strings simultaneously. The searches employ Boolean logic in one of four ways. Up to 20 text strings can be searched for in either an *O R* or an *A N D* way. Also these can be combined into *AND/OR* and *AND+OR* searches with a total of up to 40 text strings. For example, an *AND+OR* search might look for ("pheromone" AND "Scolytidae") and ("stridulation" O R "acoustic" O R "vibration" O R "sound") in each reference. Before searching for the text strings, the user is asked whether he or she wants to save this search profile for later recall. If the search profile is saved, then the next time a search is made the user may select option five to load the search profile and begin the search without the need for entering the text strings. During the search of the database, the program displays any located references on screen and waits for keyboard input to mark (or

```
(A)
AN 37011682 8907.
AU BYERS-J-A.
TI CHEMICAL ECOLOGY OF BARK BEETLES.
SO EXPERIENTIA-BASEL 45 (3). 1989. 271-283.
IN DEP ECOL, ANIM ECOL, UNIV LUND, S-223 62 LUND, SWEDEN.
LG EN.
YR 89.
KW REVIEW     DENDROCTONUS-BREVICOMIS     ELECTROPHYSIOLOGY    PHEROMONE
   BIOSYNTHESIS   OLFACTION   PERCEPTION   ORIENTATION   CUE   ANEMOTAXIS
   INTERSPECIFIC   COMPETITION   HOST   TREE   SELECTION   ATTACK   DENSITY
   AGGREGATION BEHAVIOR SEMIOCHEMICAL ATTRACTANT NORTH AMERICA EUROPE.
BC 75304 Coleoptera.
ST ANIMALS#, INVERTEBRATES#, ARTHROPODS#, INSECTS#.


(B)
CHEMICAL ECOLOGY OF BARK BEETLES.
AU  BYERS-J-A.
JO  EXPERIENTIA-BASEL 45 (3). 1989. 271-283.
AD  DEP ECOL, ANIM ECOL, UNIV LUND, S-223 62 LUND, SWEDEN.
KW  REVIEW DENDROCTONUS-BREVICOMIS ELECTROPHYSIOLOGY PHEROMONE BIOSYNTHESIS OLFA
CTION PERCEPTION ORIENTATION CUE ANEMOTAXIS INTERSPECIFIC COMPETITION HOST TREE
SELECTION ATTACK DENSITY AGGREGATION BEHAVIOR SEMIOCHEMICAL ATTRACTANT NORTH AME
RICA EUROPE. 75304 Coleoptera.


(C)
Byers, J.A. 1989. Chemical ecology of bark beetles. Experientia-basel 45 (3).
271-283.==review dendroctonus-brevicomis electrophysiology pheromone
biosynthesis olfaction perception orientation cue anemotaxis interspecific
competition host tree selection attack density aggregation behavior
semiochemical attractant north america europe. 75304 coleoptera
```
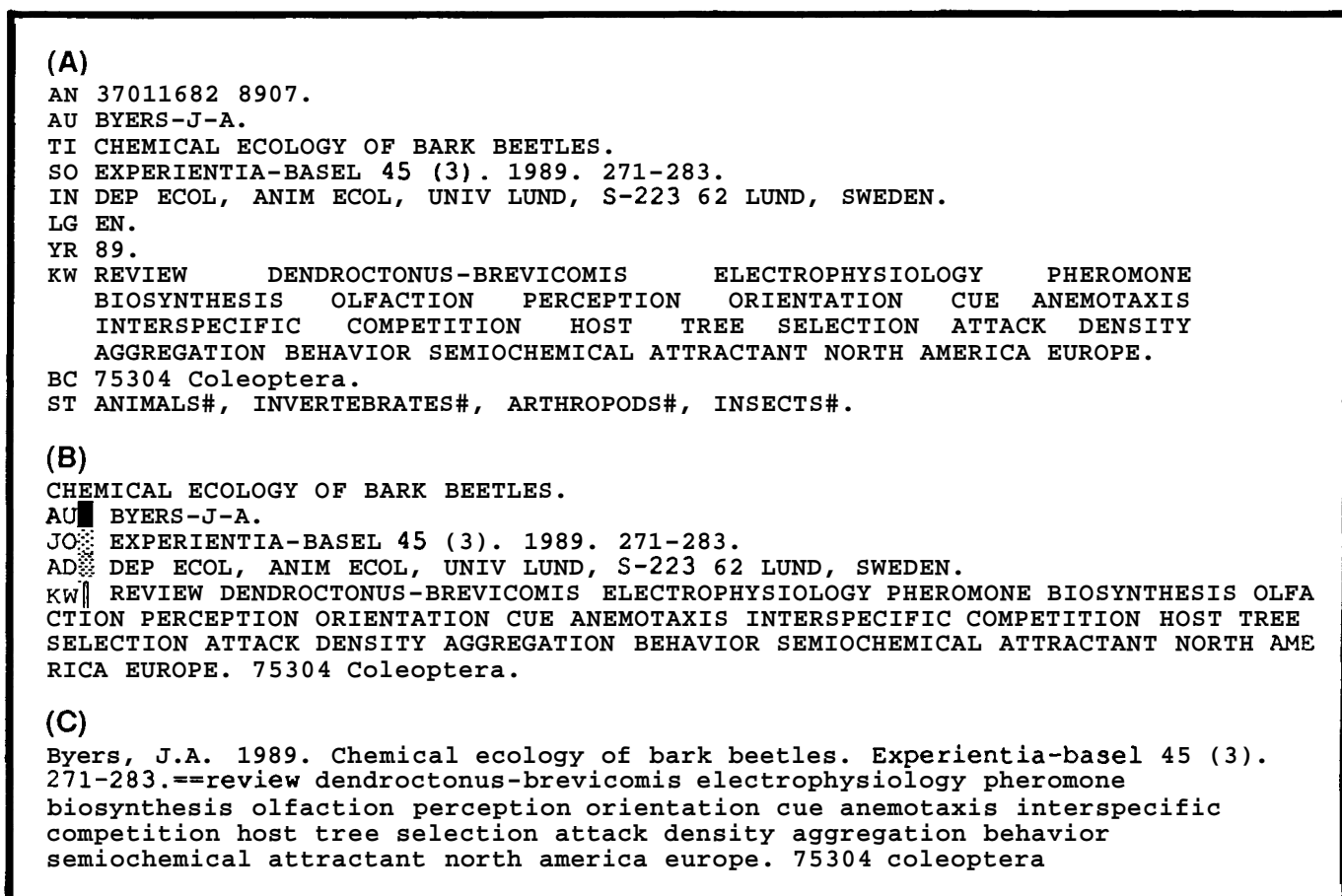
Fig. 2. BIOSIS.EXE processes a BIOSIS text file (A) and converts it to a database style (B) which can be formatted to the standard scientific style (C) or Pro-Cite style (not shown).

unmark) or continue without marking until all references in the database are searched. A search record showing the number of references found out of the total number in the database is displayed at the end of the search.

The fourth option, "Make Reprint Request Cards of Marked References in Database," asks for the user's name and address, which is saved in the file **BIOSIS.ADD.** Afterward, the choice between whether the printing will be on paper (output in Fig. 3) or modified for use on Request-a-Print cards (Institute for Scientific Information) is made. The program then goes through the database and formats the marked references into reprint request cards using the user's address and the form option. The "cards" are saved in a file called PRINTOUT, which is overwritten each time this facility is used. This file can be printed directly on perforated Request-a-Print cards or on perforated computer paper. At the DOS prompt simply type copy printout *prn,* and the file will be printed. For laser printers, the file PRINTOUT should be imported into a word processor as DOS text and the "cards" arranged to fit on the pages. An IBM graphic printer emulation yields pleasing printer output from a word processor.

The fifth option, "Format Marked Refs. in Database to Standard Refs. in Text File," is used to take the title, authors, author address, journal, and key word strings and to reformat them into a standard scientific citation style (Fig 2C). These are saved in a DOS text file called OUTPUT, which is overwritten each time this option is chosen. This file can then be imported into word processing programs as well as into the scientific reference manager, CardCat (Byers 1991).

Considerable programming was necessary to convert the several BIOSIS punctuation styles for authors into the standard citation form. The date (1900s or 2000s) is extracted from the journal string. The title and journal letters are uncapitalized and then the first letter of each word in the journal is capitalized. This saves time in the word processor as most letters need to be lowercase in a standard citation. An additional feature is that marked references can also be exported in a format compatible with MacIntosh computers for use with the Pro-Cite (Personal Bibliographic Software, Ann Arbor, MI) bibliographic manager software. In this case, the MacIntosh system utility for text file conversion is used to convert the OUTPUT file (which now includes a Pro-Cite code,

tab stops, and various slashes) from IBM to MacIntosh format. This file is then imported into Pro-Cite in the short journal form. Of course the use of Pro-Cite on IBM-compatible computers does not require the conversion step.

A sixth option includes either a global unmarking or global marking of all references in the database. The "Change Databases" option allows for an unlimited number of databases. Each time the program is started it gets the current database, which is permanently stored in the file **BIOSIS.DBD** until changed by the user. The user can receive help from within the program by reading the file **BIOSIS.TXT,** which includes the contents of this article with figures. This file appears to be a standard sequential DOS text file but is actually a random access file which allows scrolling the text forward and backward on the screen.

Deitz & Osegueda (1989) discuss the effectiveness of retrieving entomological literature from several databases, including AGRICOLA, AGRIS International, BIOSIS Previews, CAB Abstracts, Dissertation Abstracts Online, Life Sciences Collection (Cambridge Scientific Abstracts), SCISEARCH, and Zoological Record Online. They show that pertinent literature for a particular subject is likely to be found in any of these databases in **varying** proportions and redundancy. While the cost effectiveness may vary depending on the blend of databases used, it was apparent that BIOSIS Previews had the most relevant references.

I have searched CA Search and BIOSIS Previews for all references in the Journal of Chemical Ecology and found that the former contained only about 86% of the articles. whereas BIOSIS Previews had all of them. Several key word searches and reference retrieval strategies have been proposed (Harvey 1979, Marshall 1979, Chiang & Chiang 1984, Eyers & Taylor 1988, Resh 1988, Deitz & Osegueda 1989). The results of these articles indicate that an exhaustive search of the literature is more costly and difficult than one might anticipate because of the scattering of references and the complexity of designing a key word search profile. However, a search of the largest database (BIOSIS) with the appropriate key word profile will retrieve the majority of the pertinent references.

The advantage of building one or more personal databases from a commercial database such as BIOSIS is that the references in the personal databases can be searched for many combinations

Dear Dr. Byers:	Requestor's name Dept. of Entomology 125 Everett St. Lakewood, Colo. 80226 USA	Dr. Byers DEP ECOL ANIM ECOL UNIV LUND S-223 62 LUND SWEDEN
Please send me a reprint of your article in: EXPERIENTIA-BASEL 45 (3). 1989. 271-283		
Entitled: CHEMICAL ECOLOGY OF BARK BEETLES... Thank You,	.... use mailing label above Requestor' s name Dept. of Entomology 125 Everett St. Lakewood, Colo. 80226 USA	.... Card sent: 11-08-1991 Dr. Byers CHEMICAL ECOLOGY OF BARK BEETLES... EXPERIENTIA-BASEL 45 (3). 1989. 271-283
Requestor' s name		keep for your records

Fig. 3. Reprint requesr card (in rhe text file: PRINTOUT) formarred from rhe BIOSIS reference in Fig. 2 **hy** rhe program BIOSIS.EXE. The person requesring reprinrs deraches rhe upper righr address and glues ir to rhe opposire side for use in mailing. The lower righr address is also derached for use as a record of rhe request. Upon receipt of rhe card, rhe aurhor detaches rhe address in rhe middle secrion (now rop righr) and arraches this ro an envelop conraining the requesred materials.

of key words. This is usually less costly, and more convenient than to search the commercial databases multiple times.

Users should be aware of the 'fair use and copyright' issues concerning the use of databases. Many database companies have the policy that downloaded information may be freely used without payment when temporarily stored in lieu of hard copy or printout form. However, the temporary storage/use ceases when a user alters the archival characteristic and incorporates information into an active system of any type. In this case, payment is often required. BIOSIS, for example, charges about 15 cents per reference per year when one has more than 500 references in an active system.

The BIOSIS.EXE program system can be used to condition BIOSIS Previews files before importing them into the scientific reference manager CARDCAT (Byers 1991). Another system, JCE-REFS, can convert BIOSIS Previews files of references containing abstracts into a database that can be searched (Byers 1992). Any of these program systems can be obtained from the author by sending a formatted disk to the author. Donations of $5 or 35 Swedish Kronor would be appreciated for shipping of each program system requested.

Acknowledgments

References Cited

Byers, J. A. 1991. Cardcat: A scientific reference manager for IBM-compatible personal computers. Am. Entomol. 37: 240–242.

Byers, J. A. 1992. Trends in chemical ecology revealed with a personal computer program for searching data bases of scientific references and abstracts. J. Chem. Ecol. 18: 1481–1495.

Chiang, K. S. & H. C. Chiang. 1984. Literature retrieval in entomology. J. Entomol. Res. 8: 86–88.

Deitz, L. L. & L. M. Osegueda. 1989. Effectiveness of bibliographic data bases for retrieving entomological literature: a lesson based on the Membracoidea (Homoptera).Bull. Entomol. Soc. Am. 35: 33–39.

Eyers, J. E. & E. R. Taylor. 1988. Online searching of parasitology literature. Parasitol.Today 4: 324–328.

Harvey, S. 1979. CAB/CAIN evaluation project: a comparative study on the performance of two agricultural databases in a computerized current awareness service. (British Library Research and Development Report 5483).Centre for Agricultural Publishing and Documentation, Wageningen,The Netherlands.

Marshall, K. E. 1979. Online retrieval of information: a comparison of different systems used to produce a bibliography on Ephemeroptera and pollution, 1969–78. Canada Fisheries and Marine Services Technical Report 878, Winnipeg, Manitoba.

Resh, V. H. 1988. Publication patterns in entomology: an example based on aquatic insects. Bull. Entomol. Soc. Am. 34: 145–150.

❏

John Byers has a Ph.D. in entomology from the University of California at Berkeley. He currently is a hogskolelektor (associate professor) in the Department of Ecology, Animal Ecology, and a member of the Pheromone Research Group, Lund University, S-223 62 Lund, Sweden. His interests include chemical ecology of bark beetles and computer simulation in behavior and ecology.