

The BRYOLOGICAL TIMES

Newsletter of the International Association of Bryologists

ISSN 0253 - 4738

March, 1990

Computer Special

No. 54

**Proceedings of the IAB
Computer Workshop
Mont Rigi, Belgium
Feb. 24.-27. 1989**

IDENTIFICATION

PROGRAMS

by B.J. O'Shea

The computer techniques column has become a kind of a computer user group in the IAB. It is like other computer user groups dedicated to the promotion of computer applications, in this case in botany, and it is run by volunteers. It is based on the philosophy of sharing programs and information, not selling them. Today we have great technological improvements in computing but information about it, especially information on how to make use of it for science or bryology, is poor. Therefore users who are experienced in this field share their knowledge with others. This concerns not only new developments in computer techniques, but also fundamental experience, since there are always beginners in this field.

For these reasons, our first Computer Workshop took place in Belgium a year ago (see LaFarge-England & O'Shea, Bryol. Times 50:1-3, 1989). Some of the topics of the workshop which may be of wider interest are summarized here in this third computer special edited by Jan-Peter Frahm and Brian O'Shea.

CONTENTS:

Identification programs.....	1
New Entries to the softwareLibrary.....	2
Atari Software.....	6
Disk Compatibility.....	6
Programs of the Software Library: LOAN, LITKEY.....	7
Scanning and Digitizing of Illustrations.....	8
Herbarium Management.....	9
Distribution Mapping.....	10

The traditional method of botanical identification using dichotomous keys is fairly easy to implement on a computer, but serves no point, being too inflexible, too long-winded, and possibly more difficult to use than an existing printed key.

Many polythetic keys are now published (matching on more than one attribute at a time) and these provided, a few years ago, an early use for punched cards, especially edge punched cards. As soon as it becomes necessary to compare on a number of attributes simultaneously, it is then possible to use the power of the computer in a constructive way - a way that provides new tools to the taxonomist that would not be feasible without a computer.

The field of expert systems in computing has attracted attention to those working in the field of identification, since the intention of an expert system is to emulate the expertise of an expert, and botanical identification is a field where expertise is manifestly difficult to acquire, and in fairly short supply. Most expert systems are rule based, that is, they contain a number of rules that the program applies to the data it is given. Effectively, botanical identification keys have worked in more or less this way for generations, so unless the program author is careful this approach tends to produce a traditional dichotomous key - which may be a startlingly new way to look at some problems, but not botanical identification. Also, if implemented using a declarative language e.g. Prolog, it will run very slowly because each taxon will need at least one rule to identify it - which could be thousands for the flora of one country.

We have four examples of identification programs in the IAB Software Library:

#17 Bids

#40 Swiss bryophyte key

#41 Pankey

#51 Askataxa

I am unfamiliar with the Swiss bryophyte key, so the following paragraphs refer only to the other three.

In practice, these identification programs look fairly similar in the way they work - and this is a case of convergence rather than plagiarism: there is clearly an obvious way to tackle botanical identification. They appear to operate by providing a simple two dimensional table of data (taxa being one dimension, attributes another), with the cells in this matrix containing data of a complexity dependent on the sophistication of the program.

Although they differ in detail, in presentation, and in the distance they have progressed, they can be shown to have (or to intend developing) the following attributes:

- characters can be input in any order

Rather than requiring the identifier to eliminate all the obviously unusual taxa first (e.g. Sphagnum, Andreaea), the system allows the available characters to be given in any order. This avoids those difficult places in keys where it is impossible to progress without a piece of information that is absent in the specimen you have, such as peristome characters in a non-fruitlet specimen, and also provides an easy route into the identification for beginners, by using characters that are easy to recognise.

- computer to either calculate probabilities for each taxon as a result of receiving input, or do direct elimination of taxa incompatible with input data

Each piece of information gathered

continued on p. 6

NEW ENTRIES TO THE IAB

SOFTWARE LIBRARY

Since the last announcement of new entries to the software library in the second computer techniques supplement, 115 diskettes have been added. At present, 238 diskettes are offered free for IAB members covering all possible applications for IBM and compatible computers. The IAB software library is so far the only scientific software library. It has been offered for the use of lichenologists in the International Lichenological Newsletter, and for taxonomists in general in TAXON, Delta Newsletter and Herbarium News. More than 1000 copies of the programs on diskettes 1-122 were distributed in total. The new additions have partly been made available by colleagues, who freely provided their programs. Numerous contributions were made by J. M. Glime and A. Empain. R.H. Zander added (amongst other programs) all FLORA ONLINE volumes with bryological contents to the software library. Other programs have been added by exchange and a large amount by purchase. The copies which have been so far distributed free have a value of about 5000 \$ if bought from the public domain software market, and many of our programs are not even available from this source. To cover the costs for buying programs and maintenance of the software library, a small fee of \$5 per diskette will therefore be charged for institutions and nonmembers of the IAB. Members of the IAB will continue to get the programs free. However, if possible, you may support this unique service of free computer programs by sending an institutional order.

#123 Hypergraph (A.M. Empain)

A demonstration of an on-line graphic environment for taxonomy. Includes descriptions, references and type citations with windowing and illustrations of species and distribution maps that allow zooming for more or less detail, including a coloured photomicrograph if you have an EGA screen.

#124 Archivist (IRL Press)

Demo disk (maximum 16 entries) of a database manager and text processor for bibliographic

reference handling. This is a professional program aimed at scientists, and includes the ability to store and access references, including searches by keyword, as well as being able to format output according to templates for different journals. Price of full version: £ 47 + VAT.

#125 COENOS (A. Ceska)

Test version of a program for identifying species relevée groups in vegetation studies (see Ceska, A. & H. Roemer, Vegetatio 23: 255-277, 1971). This version can run to 50 relevées and 200 species which is sufficient for smaller, but still reasonable data sets.

#126 EDRAW v 3.2 (shareware, PC-SIG 828)

Easy Draw drawing program with mouse support for IBM/Epson and HPLJ printers. CGA or EGA required.

#127 PC-LOCK 1.0 (freeware for non-commercial use)

Password protection for your hard disk or individual directories, encoding-decoding and hiding-unhiding of specified files.

#128 Mandelbrot Magic (shareware, PC-SIG 841)

Program to create fractal slides of images of many different subjects; with demonstration program. Requires CGA.

#129 PC-DEMO (shareware, PC-SIG 914)

Graphics presentation package that can create demonstrations featuring graphics animation and slide shows.

#130 Chemical Molecular Modelling (shareware, PC-SIG 938)

Forms three dimensional models of chemical molecules. The user can pick up atoms from the Periodic Table which can be bound into molecules, or molecules can be bonded with others. The display can be rotated and seen from different angles. Requires 640 K RAM and CGA, and EGA for some of the programs.

#131 Flop Cat (PC-SIG 959)

Disk cataloguing system, menu driven, with disk label feature. Reads disk information, searches for files, prints listing of disk contents. Requires two disk drives.

#132 Weak Link (shareware, PC-SIG 893)

A networking program that connects two PCs via serial ports. Allows to transfer files between two machines which have different disk formats (like a PC with 5.25" drive and a laptop with 3.5" drive). When installed, the master unit can address the disk drives of the second machine with D:, E: and F: and transfer files with a speed to 115 Kbaud. Requires cable or adapter with crossed wires for transmitting and receiving data. Do-it-yourself instructions for modifying a normal RS232 cable are provided on disk. This program comes with Private Line, an encryption/decryption program for data files.

#133 IMAGE 3D (shareware, PC-SIG 762)

Same program as #113, but for CGA.

#134 Disk Cataloguing Program (shareware, PC-SIG 1049)

Reads disk information, changes disk labels, lists disk contents, searches for file-names. Individual file descriptions can be added once the disk is catalogued.

#135 Bradford (shareware, PC-SIG 1053)

Prints any ASCII or Wordstar file in different font on dot-matrix printer. Prints Greek, Russian, Hebrew, double height, typewriter quality, proportional spaced, or italics. Over forty fonts are included; additional fonts can be created.

#136 Express Graph (shareware, PC-SIG 1058)

Creates line graphs, bar graphs (overlapped, stacked, 3D), and pie charts (uncut, separated) in a variety of styles and formats. Data can be entered or imported. Graphics can be flipped, combined with text, lines and patterns. CGA or emulation required. Graphics.com must be loaded to get printed output.

#137 Form Master (shareware, PC-SIG 1099)

Generates, modifies and prints forms. Form Master supports most printers with features including boldface, italics, underline, expanded and compressed text and can export files into other software.

#138 CENOSIS-2 (E. Ezcurra & M. Equihua)

Multivariate Analysis for ecological and floristical studies, especially for classification of vegetation units. With Pascal source code. Dialogue in Spanish.

#139-140 Bryoflore des Vosges (J.-P. Frahm) Bryoflora of the Vosges-Mountains and surrounding areas. Complete manuscript in French (400 K). Available as ASCII or WORD-file (please indicate).

#141-147 COMMANAL (J. Glime) Extensive program package for ecological calculations. The programs provide structured data entry and checking for reasonableness of the data. They permit separating data by zones or ranges of resources. They compute mean cover and frequency. They graph mean cover in each up to 10 resource states by species for each resource. They compute Shannon and Brillouin species diversity by sample or location. They compute Levins and Freeman-Tukey niche width and overlap and graph the results. Output is in ASCII in the form of formatted tables, graphs, and tables suitable for loading in a spreadsheet or statistical program. Documentation on disk. Updated version of disks # 28-35.

#148 Herbarium Label Variations Modifications of the herbarium label program LAB on disk #2:
LABCOND: with header and lines 60 characters wide for condensed print.
LABALTA: with 2 lines header and 12 lines of 40 characters.
LABNEW: without header and with 11 lines 60 characters wide.
LAB1140: as labnew, but lines 40 characters wide.
LABEXP: 8 lines with 3 lines header, 40 characters wide, header expanded.
All programs with source code in BASIC.

#149-151 INDEX (J. Glime) Program for creating an index. Requires ASCII file with page markers.

#152 KFYOUT 2/2 (J. Glime) Second disk for # 121.

#153 LANGUAGE V 1.0 (B.J.O'Shea) This programs allows you to enter literals for a program in any language to create native language versions.

#154-155 Herbarium Label and Record Book Program 2/3 and 3/3 (J. Glime) cont. from #3. Contains 150 K species list with authors for bryophytes of North America (154) and 190K species list for Japan, Iceland and Europe (155).

#156 TREESEARCH (Bruggeman et al.) Cladistic program written in Turbo-Pascal developed at the university of Utrecht, The Netherlands. Requires no math-coprocessor.

#157 Picture Library (J.-P. Frahm) Scanned illustrations of bryological subjects for use in Paint-programs, slide shows, teaching, pasting into text or desktop-publishing. Contains .pcx (PC-Paintbrush) and .img (GEM-Paint) files. Still only a demonstration for possible applications.

#158 Bibliographer (IRL Press) Demo-disk of a literature database program. The full version is available for US\$ 99.

#159 Big Softy (AnyTime Software) A data base of Public Domain and shareware software, which can also be used to built up a personal data base. Is nicely menu driven with windows and can search for subjects.

#160 TWINSPAN, DECORANA (McConnell University) Two way INDicator Species ANalysis. A program for arranging multivariate data in an ordered two way table by classification of the individuals and attributes. The two way classification makes possible a tabular matrix arrangement which approximates the result of Braun-Blanquet tablework.
DEtrended CORrespondence ANalysis, a program designed to make ordinations of data on the occurrence of a set of species in a set of samples.
Version of these programs which do not require a math-coprocessor.

161 MP-SORT (B.M.Möseler) Program for writing and automatically sorting plant sociological tables with advanced features such as calculation of frequency etc. Requires MS-Multiplan.

#162 INFO Advanced dBase III plus program for filing literature. Supports several databases, has

comprehensive search functions in all fields, has help functions, a window to DOS and output as filecard or ASCII-file. Dialogue sin German. See also #169 and #234.

#163 ECOTROP Bibliographical database on the ecology of tropical bryophytes. Requires dBase III+ and #162 Info.

#164 Computerized taxonomic descriptions (A.R. Empain) Encoding and Decoding of taxonomic descriptions including automatic translations and an interface to Wordperfect.

#165-166 PC-Calc+ (Shareware, PC-SIG #199, 1016) Spreadsheet program, advanced version, now allowing the user to import Lotus 1-2-3 files and to create graphics.

#167 MEKA V. 1.2, MEKAEDIT V. 1.0 (Th. Duncan & C.A. Meacham) A Multiple Entry Key Algorithm that allows identification of an unknown specimen. Mekae-dit allows users to make their own keys for access by Meka.

#168 Bookworm (shareware) The only bibliographic database manager running under MS-Windows. Simple to use because it is based on the MS-Windows graphic user interface. Has full ANSI and ASCII character set support (incl. all foreign characters), copies and pastes references into a manuscript using Windows clipboard, has custom formatted output according to specific journal specifications, imports references from textfiles, and searches for 2 or more whole words in entire database or selected range. Requires MS-Windows, hard drive, and mouse. Still prerelease version.

#169 INFO Enhanced version of program #162 with sort/browse/delete functions and modified database structure. See also #234.

#170 MYSTAT An interactive statistics package with a large number of commands to analyze data (regression analysis, correlation, t-tests, cross tabulation, analysis of vicariance etc.). Data entry by some kind of spreadsheet. Includes graphical

Output as plot diagram or histogram. This program may not be distributed but single copies are available free, with printed documentation, to those who write to SYSTAT Inc., 1800 Sherman Ave., Evanston, IL 60201, USA.

#171 HERBAR 2

English version of the program on disk # 86. Comes with fixed fieldlengths and a special screenform for the label. Supports two different label formats, which can be modified.

#172 HERBAR 3

Modified version of program # 171. Has no fixed screenform, allowing users to modify the structure of the database to their own requirements and thus to adapt it to any required herbarium label. The program allows the input of label data, print of any number of any record in two different print formats, to print and modify the last entry or to modify, replace, append and print any record. Those who want to adapt the program to their special requirements and are not able to make the necessary changes in the structure of the database and the labelform, may contact the distributor.

#173 dBASE UTILITIES

dblist (B.J.O'Shea) converts dBase III and IV files to ASCII format so that it can be used outside dBase.

#174-176 PC-FILE:db (shareware)

The classic shareware database program, now dBase compatible. Menu driven. Allows you to generate reports, labels, formletters and now even bar-, pie-, line- and scatter-graphics from a databasefile. The package includes a program for converting files generated with previous versions of PC-File. Supports up to 1 billion records per database, and up to 70 fields with 4000 characters each.

#178 PHYLIP 2

One of the programs (Mix) coming with PHYLIP (#44-49), which was compiled by R. Bayer (University of Alberta) and so does not need a Pascal compiler. Comes with a prepared data matrix and a demo output file.

#179 KERMIT (freeware)

The classical data communication program for the use with modems or also for data transfer from one computer to the other via serial ports.

#180-181 PC-Write V.2.7 (shareware in the US and Canada)

Now with mouse and laserprinter support. All versions of PC-Write produce pure ASCII files which can be read by any other wordprocessor. Can be used also to "filter" other files to pure ASCII files.

#182 ANTI-VIRUS PROGRAMS (freeware)

Antivirus programs FLUSHOT, FLUSHOT+ and CHECK4BOMB, will be installed resident and watches for any changes of command.com and other DOS programs.

#183 Archiving-Unarchiving Programs I (freeware)

Programs that can compress (and extract) contents of diskette and provide ca. 40% more storage capacity: PKX34A20, ARC600, PKX35A35, PKZ090, ARCE.

#184 DMAP II (A.Morton)

Shareware version of disk #112, a program for distribution mapping at any scale. Automatically superimposes a numbered grid if required. A range of symbol types is available, and the size of symbols is user-defined. With facility for coincidence mapping of groups of species. Runs on all commonly fitted graphics displays. Supports 24 and 9 pin dot-matrix printers, and also Hewlett-Packard-compatible ink-jet and laser printers giving high resolution publication-quality maps.

#185-188 Laserjet Fonts 1-4 (shareware)

Provides a broad variety of additional fonts (Centrum, Optimis, Garamond, Helvetica, Old English and Script) in different sizes (8,10,12,18,24 points) in normal, bold and italics. Comes with download software.

#189 Laserjet Fonts 5 (shareware)

10 and 18 point Script, TimesRoman, Oldelish,

Helvetica and Hebrew and also Symbols and Line Drawings.

#190 Label Maker

Simple program with 5 lines input without special printer codes. Also usable for identification strips.

#191 Listmate

Menu-driven label producer. Works in conjunction with Multimate for formletters.

#192 Labelmaster V. 3.0 (shareware)

For users to print mailing labels or other types of repetitious printing. Can read Lotus 1-2-3, Symphony and dBase Files.

#193 Doctor Data Label V. 3.1

One of the most powerful menudriven and flexible mailing list management programs. Has user-definable custom label formats, ability to index fields, mail merge, and is dBase compatible. Comes archived.

#194-195 As Easy As V. 3.01 (shareware)

A Lotus 1-2-3 clone, looks and runs just like it. Can import Lotus files.

#196 Fingerprint V. 2.0

New version of # 74, now supporting EGA, CGA and Hercules graphics cards and the use of a mouse.

#197 SAIL V. 2.0

A powerful texteditor, can edit several textfiles at a time in several windows.

#198 XYFIT & SURF (shareware)

XYFIT allows you to fit data to a line, a polynomial or smooth curve. Can zoom part of the curves, calculates correlation coefficients. SURF is a 3D plotting program.

#199-200 WINDOW BOSS (shareware)

Programming help in C to generate Pop-up-windows, pulldown menus, status lines, multiple windows at the same time or onlinehelp.

#201 MAPS (A.Empain)

Maps of any part of the world can be generated in different projections with or without political borders and grids, and completed with a variety with different symbols and text. High quality output, Postscript emulation compatible, with 300 dpi.

#202 Newkey (PC-SIG #181)

Keyboard utilities to redefine function keys, generate macros and to extend keyboard buffer.

#203 Easywrite (PC-SIG 146)

A simple wordprocessor, especially for use as a texteditor to manipulate batch files or readme files.

#204 Pottiaceae files (R. Zander)

Pottiaceae of North America North of Mexico; A bibliography for the Pottiaceae (Musci); ASCII files from Flora Online vol. 8.

#205 North American Mosses

Paul L. Redfearn, jr.: Checklist of the mosses of the Interior Highlands of North America in Arkansas, Illinois, Missouri and Oklahoma (Flora Online vol. 12).

J.A. Snider, S.M. Studlar & M. Medley: A list of the bryophytes of Kentucky, with county distributions (Flora Online vol. 15)

#206 PC-Taxon Utilities

MAKETAXN, MAKEDEL: convert files between the taxonomic programs PC-Taxon and Delta, with instructions for the use of each program. (Taken from Flora Online vol. 21).

#207 Baker's Dozen (Shareware, PC-SIG #800)

Utility programs from the author of PC-File + and PC-Calc +. Includes a memory resident calendar, a disk program to recover erased files, a file finder that searches strings of text, a side-ways printer, a mini-spreadsheet, a program to swap COM and LPT ports, and many others.

#208 Utilities 2

IBM HELP: a better approach to the IBM extended character set. Memory resident.

E.Q.Filer: Accepts input data for equations and solves them.

SPLIT: a program that splits files

#209 Mapmaker (PC-SIG #219)

Program designed for making computer plotted maps. It allows the user to turn statistical data based on area into a form. Comes with sets of maps for Florida and the US. Requires CGA or EGA.

#210 ICONVERT

A graphics format conversion program to convert graphics from Dr. Halo, First Publisher,

Fontasy, GEM, GIF, IFF, MacPaint, Windows-Paint, PC-Paintbrush, Printmaster, Printshop and others and vice versa. In addition, it allows previewing, flipping, inverting, clipping and mirroring of graphics.

#211 MATHPAD

A scientific calculator with highly advanced features such as trigonometric functions, linear equations, vector analysis. With 22 memories.

#212 PC Magazine Utility Disk I

Collection of 43 utility programs with documentation.

#213 PC Magazine Utility Disk II

Another disk full utility programs.

#214 Conversion programs

CONV-CONF: Qmodem conversion program
TEXTCON: ASCII file converter, removes carriage returns at line ends, removes blanks and extra lines. For importing of files into word-processors, eliminating problems.

WSPILTER: Converts Wordstar files to ASCII so that they can be typed by DOS.

#215 Encryption

Collection of two encryption programs, "The Confidant" (confide2) and CDES.

#216 Simulation programs

REGRES: Program to perform simple linear regressions on a set of data.

NNS: Neural Net Simulator. Simulates 256 electronic neurons.

GFS: Gravity Field Simulator; simulation of gravitational fields.

#217 Directory Master

A DOS helper for all file procedures (copy, rename, view etc.) and DOS commands, with facility to program function keys.

#218 Archive-Unarchive Programs II

Collection of the self extract utilities CAS, SARC, FSRCHA and PKSPFX.

#219 LOAN II (G. Wagenitz)

English Version of the Herbarium Loan Program (#103). Requires dBase III +.

#220 LOAN III

Modified Version of #219 for the Herbarium of the University of Alberta. Has some more fields

for number of parcels and exchange-balance.

#221-230 Bryological Bibliography (J. Glime et al.)

More than 13.000 (!) references as ASCII files on 10 diskettes. Can be used as textfile. For database-evaluation, the program GaSP (#235) is required.

#231 Bryophytes of North America

A dBase version of the ASCII file of North American bryophytes set up by J. Glime as spellchecker for the herbarium label program on disk # 154-55 prepared by B.J. O'Shea. The list contains some more taxa from Europe and Japan.

#232 Bryophytes of Europe

The same as #231 for Europe.

#233 Checklist of the Mosses of Canada (R.Ireland et al.)

Wordperfect and ASCII file.

#234 Info III

English version of the bibliography program on disk # 169.

#235 GaSP (J. Ridley)

Bibliography search program. Allows the user to work with the comprehensive bryological bibliography on disks nos. 221-230. With demo file and documentation. Comes with a program which converts Current Contents on disk to GaSP format.

#236 LAB V. 1.5 (B.J. O'Shea)

New version of the herbarium label program on disk # 2. Allows to print either 40 normal or 60 condensed characters per line for more information storage. Defaults are shown on screen. Now with multilingual dialogue. New languages can be added using the language program (#153).

#237 VIRUS DIAGNOSIS (shareware)

SCAN V.48: Identifies virus infection, indicates infected files and identifies the type of virus. Scans (hard)disk for 48 different viruses, including the most common 10 which account for more than 95% of all infections. Comes with a program to validate the integrity of scan.exe.
M-VIENNA: Locates and repairs (!) files infected with the Vienna virus.

NETSCAN V.46: The network version of

SCAN. Scans network virtual drives for infection with 46 viruses.

The complete catalogue of the IAB software library is available on diskette. It is menu driven and allows to read the list of programs, a subject and an alphabetical index, and also to print the catalogue and order forms. For a free catalogue, send a blank 5,25" disk to Jan-Peter Frahm, Universität Duisburg, Fachbereich 6, Botanik, Postfach 101503, 4100 Duisburg, F. R. GERMANY.

NEW - NEW - NEW - NEW - SOFTWARE LIBRARY FOR ATARI ST COMPUTERS

(G) Programs in German

#1 Graphics programs (G)

ER-CHART: Business graphics

GFA-BILD: Graphics manipulation

#2 Disk & Printer Utilities (G)

SAGROTAN: Anti-virus program

PROTECT: Write protect for harddisks

UMLAUT: Printer adaptation for special characters

#3 Database programs (G)

P_A_D_M: dBase type of database program

Literatur: Bibliographic program

GEM_CALC: Calculation program

#4 Various programs

UNITERM v. 2.0a: Terminal program

UTILITIES: Ramdisk, fast mouse driver, file archiver etc.

#5 Wordprocessors (G)

Three different programs

This new software library will be extended if there is any interest in Atari programs. Contributions to this new software library are welcome. As usually, all programs are available free for IAB members by sending the appropriate number of blank unformatted, in this case 3.5" disks to J.-P. Frahm.

These programs have been provided through the generosity of colleagues who are pleased to be able to share their knowledge, and to make their efforts freely available. If you find any value in the programs you receive, please think about any contributions of public domain programs or data you might make, or any suggestions you can give of where such programs could be sought to add to the IAB Software Library.

during an identification session reduces the remaining search space. The program can respond directly by eliminating from further consideration those taxa that do not possess the attribute the identifier has described, or can quantify the information using probabilities that are accumulated during the session. Those that are to be 'eliminated' are then only decided on at the end of the session, by virtue of each taxon's overall likelihood of being the identity of the specimen being examined. The specimen's identity is likely to be amongst the taxa with the highest probabilities.

* computer maintains list of likely candidates

During a session, the program maintains - and displays either permanently or on request - a list of the taxa that are the likeliest identity of the specimen. This becomes of greater assistance as the session progresses, as it allows the direction of the session to be more closely focused on the most likely identity of the specimen.

* suggests characters that will prove most useful

Based on the information already received, the program can look at which data would give it the greatest advantage in progressing the session, and thus indicate which character would provide the greatest discrimination in eliminating taxa or altering their relative probability.

* suggest characters to work towards specific solution

If the identifier has a good idea of what the taxon might be, the program can be asked what information it would need to support the hypothesis.

* the identifier must be able to change his or her mind

As we all know, it is possible using a

dichotomous key to find oneself in an obviously inappropriate part of the key, having made a wrong choice. There thus must be a means of revising information already given - which might have a dramatic effect on the course of the session. This may mean some rather messy programming if cumulative probabilities are being used.

* must allow 'fuzziness' e.g. ranges/combinations of characters

The sophistication with which this factor is handled varies widely, but inevitably means asking the user of the system for more detailed information, such as asking for the position on a qualitative scale, or giving ranges of sizes. All of these result in quantification within the program, but the vaguer the description, the less use it is.

Increasing attention is now placed on the ease of use of programs, and in each of the three IAB Software Library programs mentioned above, a great deal of effort has been spent in trying to get this right. They all look different, but work by dividing the screen into windows, with particular pieces of information always being in the same place.

Problems with disk compatibility?

There are two different disk formats in use for IBM and compatible computers, the classic 5.25" and the newer 3.5" format, which is becoming increasingly common. On an PC/XT, the 5.25" disk is formatted with 360 K and the 3.5" disk with 720 K. Since the introduction of AT computers with high density (HD) disk drives, the 5.25" disk can also be formatted with 1.2 MB and the 3.5" disk with 1.44 MB. All these different disk formats are upwards compatible, that means a high density disk drive reads normal disks without any problems.

Problems can sometimes be caused by reading

and overwriting a disk, for example, when an XT user and an AT user exchange disks. These problems concern only 5.25" disks. There are no problems with the exchange of 3.25" disks: if a XT user sends a 720 K formatted disk to a colleague with an AT and a high density disk drive, the disk drive recognizes the format, and writes corrections or changes of a manuscript in 720 K format so that the XT user will not have any problems reading the files written with the HD disk drive. This is different with 5.25" disks. If an XT user sends a 360 K disk to a colleague and the latter changes files on this disk with his HD disk drive and sends this disk back, the XT user will in most cases not be able to read this file again. What had happened? The XT user has formatted the disk drive with 40 tracks in his disk drive which has broad read/write heads. So his disk has 40 broad tracks. The AT user has a small read-write head in his HD disk drive. This disk drive can read the bits on the disk, but when it writes a file back on disk with the narrow disk head, it is written only in small tracks having only half the width of the XT tracks. If the XT user tries to read this disk again, his broad disk head reads the wide track which consists now of the narrow track written by the AT and also the junk which remains on the wide track from the original file written by the XT. Result: the file is no longer readable by the XT. To allow the exchange of disks between XT and AT users, the following must be considered: The broad XT disk drive head can read 360 K/40 tracks disks written by the small head of an AT disk drive if there are bytes only on the small tracks and no left over bytes previously written by the broad head of the XT disk drive. If, therefore, an AT user gets a disk from an XT user for modification of a file, he has to load this file on his harddisk, modify it and then reformat the old disk (or format a new disk) in 360 K format and load the corrected file(s) from his harddisk back to this newly formatted disk.

LAB SOFTWARE LIBRARY programs are now also available on request on 5.25" 1.2 MB, 3.5" 720 K and 3.5" 1.44 MB diskettes. In this case, the diskettes must be formatted and the format indicated. If you do not prefer to have each program on a separate disk, you may send one 1.2 MB disk for 3, one 720 K disk for 2 or one 1.44 MB disk for four numbers of the software library. Copies on standard 5.25" disks are free; for copies on other disk formats, a fee of \$3 is charged per disk.

PROGRAMS FROM THE SOFTWARE LIBRARY

#103, 219, 220 LOAN (G. Wagenitz)

A dBase program for the management of herbarium loans, developed by the curator of the herbarium of the university of Göttingen and therefore a very practical approach for this botanical computer application. An alternative to this program is the Herbarium Manager by Barbara Thiers, which is in use at the New York Botanical Garden. The latter is a larger program, which can only be run on a hard disk. The original version of LOAN (#103) has a German dialogue, however, for the use in the software library, the dialogue has been translated into English (#219, 220). The features of the loan-program are:

1. It writes loan forms and automatically calculates the total number of specimens. For this purpose the print form must be modified by inserting the header and all other remarks which shall appear on the loan form.
2. It inserts automatically herbarium addresses in the loan form by typing the herbarium abbreviation. The address file contains the addresses of 80 of the major herbaria in the world and can be updated immediately if an address is not present.
3. It checks to whom material of a certain genus has been sent on loan. In this way it is possible to get a quick overview of who is working on a certain genus.
4. It checks which loans are due when the loan period is over.
5. It allows all loans to be checked on screen.
6. It lists all loans from the same herbarium.
7. It writes back loans which are returned.

#96 LITKEY (G. Wagenitz)

A bibliographic program which has been developed for the use of botanists for storage and retrieval of literature (especially on the field of Systematic Botany). The major special feature is an automatic abbreviation of family-names and key-words (saving space on the disk and allowing quick input without writing names in full or looking for an abbreviation). Another feature of interest is the testing of abbreviations of journals to see if they are correct.

After the input of the usual data for a reference (author, year, title, citation) you can write the names of several genera (if appropriate). Then

you are asked, if you want to check the abbreviation of the journal. If 'Y' or 'y' for Yes is entered the citation is compared with a list of abbreviations based on B-P-H. At the moment this is only a small sample, but the list is augmented every time you use it and enter an abbreviation, which must of course be checked by you for the first time. The only thing to observe is that you are consistent in not typing a space after a punctuation.

In the next step you can give the names of up to 5 families. If there is no family-name or you want to finish, put in - or RETURN: You need not type the family-name in full, the first five to six letters are usually sufficient. They must only be unequivocal. E.g. 'Polyg' is not enough, as it could mean Polygalaceae or Polygonaceae. In this case you are asked to give the name once more, preferably in full as there is no further check. At the moment most families of the angiosperms are already in the file with their abbreviation according to the list published in Taxon 31: 75-81. 1982.

The program has already a file of several hundred English key-words, comprising many morphological, anatomical, geographical and general terms, e.g. leaf, flower, xylem, USA, nomenclature, cytology. The procedure is almost the same as with the family-names. If you use a key-word not yet in the list, you are asked to choose an abbreviation (three to five letters, signs like + or * are also possible). Please begin with a capital, the other ones may be small ones or signs). If this abbreviation is already in use, part of the list of abbreviations is shown to you in order to make an appropriate choice of a new one.

Options 1 and 3 are more or less self-explanatory, please learn by using them.

Option 4 allows the search according to a family, a genus and a key-word or combinations of two of these. Family-names and key-words are automatically abbreviated as in option 2. Another possibility is to ask for all references marked before with an *. At the same time you have to decide about the mode of output. There is the possibility to see the references on the screen, to have them in addition on a file (which can be afterwards revised by a text-editor) or to have them printed in one of the two formats A and B. You have only to choose the appropriate number from a matrix shown to you to get the search-items and the output-modes as you want it.

Option 5 gives the possibility to get a print-out in format B of all references after a given number. If you want you will thus have a complete hard-copy available.

Option 6 deletes the marking *, e.g. after a reference-list for a publication is ready.

(Text taken from the documentation on disk).

Scanning and Digitizing of Illustrations

by J.-P. Frahm

Computerized pictures can be used for a variety of different purposes. Text can be merged with pictures (especially with desktop publishing programs) to provide illustrated documents. Slide shows can be prepared for demonstrations. Illustrations or herbarium labels can be stored for documentation or archiving. Illustrations can be made usable for other purposes by erasing, adding or changing details or adding text. Illustrations will have importance in the future for computer-aided teaching.

Pictures can be digitized into a computer by following methods:

- Handy Scanner

A Handy Scanner looks more or less like a large 'mouse'. It is moved by hand over an illustration (but distortions can be caused by different speeds). The scan width is limited to 6.4 cm (10.5 cm in some models). The software supplied with a Handy Scanner usually allows the user to combine several scanned stripes of a larger illustration, but this can be circumstantial and does not give the best results. Thus the use is rather limited for larger originals, but on the other hand ideal for small originals such as herbarium labels, which can be scanned during a herbarium visit.

Several models now support up to 8 grey scales. The quality is, however, high with a resolution of 200 to 400 dpi (dots per inches). Models with 400 dpi have even a higher resolution than a laser-printer. It is even possible to scan pictures into a paint-program such as GEM-Paint or PC-Paintbrush. There the scanned pictures can be modified e.g. by adding text. The price is between 300 and 600 \$.

- Scanner adapter to a printer

Here the print head of a dot matrix printer is replaced by a scanning adapter. Sheets of paper with illustrations can be fed through the printer and line by line the image is scanned. The resolution is as high as in a Handy Scanner. The price is only \$ 250, but a special dot matrix printer (some Epson and Star models) are required.

- Flat bed scanner

This method is by far the best and at \$ 1500 on also the most expensive. As in a telefax an illustration of A4 size is digitized. Flat-bed-scanner support from 16 to 256 grey scales with high resolution.

To give an idea of the necessary resolution, here is an example: to scan a line of 1 mm width, 200 dpi are necessary. Finer lines need higher resolution. So the standard of 300 dpi in scanner fulfills most needs.

As well as illustrations, text also can be scanned. The text exists, however, only as a picture file and is not recognized by the computer as text. Special text recognition programs translate such text into ASCII codes. The text-picture is separated into horizontal lines and vertical distances between characters. By special algorithms the characters thus separated are identified. The result depends on the program used, and on the font. Simple courier fonts may have a recognition rate of up to 98%. This means, however, that in every line of 80 character there are 2 mistakes or unrecognized characters. An identification rate of 80% is even less tolerable, meaning 20 errors or unidentified characters per line. So text recognition is often not much really practicable but an interesting demonstration of computer capabilities. Unidentified characters may even be learned by the program. Proportional fonts have a much lower recognition rate. Text with mixed fonts (e.g. Times normal and italic) cannot be recognized - as cannot no handwriting, of course.

- Digitized video-pictures

Also video signals of a video camera or a video-recorder can be digitized. It requires a special interface. However, the resolution is much lower compared with a scanner and cannot be higher than the resolution of the monitor screen which is limited by the graphics card used. Resolutions of video-digitizer vary between 256x256 and 720x348 pixels according to the price range. It should not be lower than 640x400 or 512x512

pixels to get adequate results. (Adequate means a quality about as bad as a picture in a newspaper). The interface costs from \$ 300 upwards, and a video-camera is also required. The hardware interfaces differ not only in the price and the resolution but also the time it takes to digitize a picture. It varies between the instant catching of an image to 1-20 seconds.

All digitized pictures require a lot of memory to store, depending on the resolution. In the following, the storage capacity in Kb is given for an A4 page scanned black and white in different resolutions:

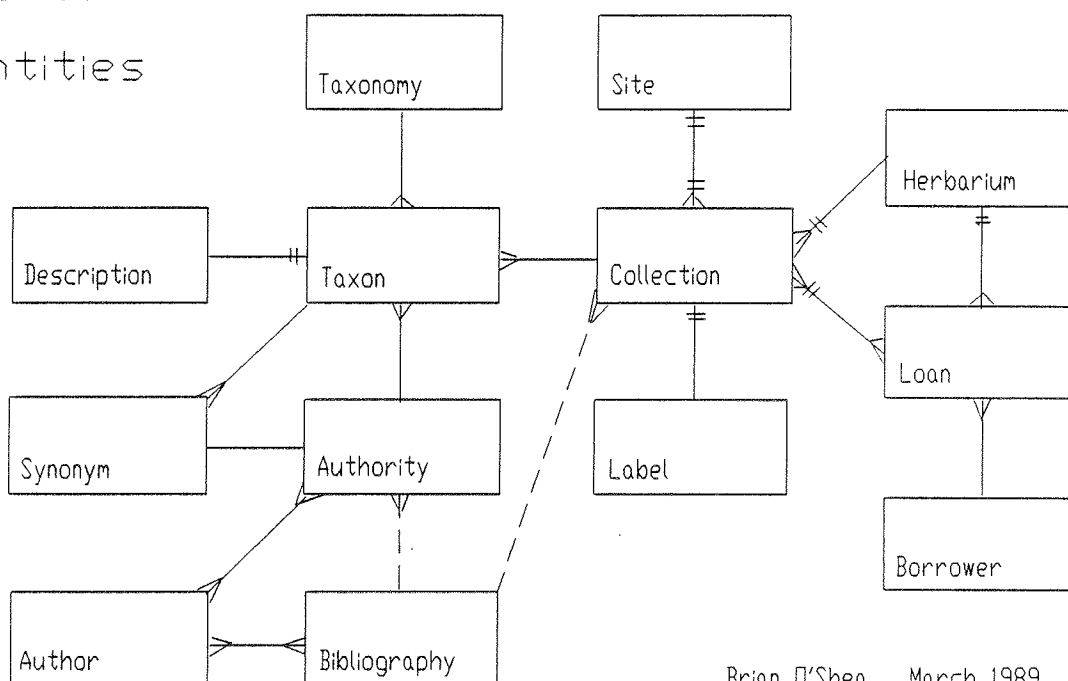
resolution	Memory Kb
200	54
400	110
500	142
600	173
700	199
800	230
1000	293

Using 256 grey scales at highest resolution, the memory requirement can be up to 16 MB per page. Thus digitizing is no matter for floppy disks.

Special attention has to be paid to the format in which a digitized picture is stored for the use with Paint- and DTP programs, e.g. .PCX for PC-Paintbrush, .IMG for GEM-Paint, or .TIF for Paint Show Plus.



Oilbodies of *Jungermannia caespiticia*, scanned with a videocamera

Herbarium
entities

Brian O'Shea March 1989

HERBARIUM MANAGEMENT

by B.J. O'Shea

As much of herbarium work is fairly undemanding and repetitive and follows standard guidelines, it seems an ideal place for computer assistance to be applied. Without a computer, it is easy enough to add new specimens to a collection, including any administrative work (e.g. card index) that may also be involved; what is then very difficult without a computer is extracting information from the herbarium about sites, collectors, taxa and so on. Often it is also difficult to keep abreast of nomenclatural changes - for instance the revision in 1981 of the names of British taxa. A computer can give assistance in both of these areas.

Even so, the management of a small herbarium poses few administrative problems, and it may be difficult to see the value of holding details of such a herbarium on a computer, rather than, say, in a card index. On the other hand, for a larger herbarium, where such a computer system may be of great value, the initial task of entering information about the existing herbarium into a computer may be immense, and the amount of data that would need to be held may

take it out of the storage capabilities of a microcomputer-based system, let alone the need for multiple access. (The capacity of a micro-based system might realistically be limited to say 100,000 accessions, depending on the amount of collection data that was stored, but it may be unrealistic to expect only one person to be able to access the system at any one time.)

The motivation for producing a herbarium system of my own (BACCHUS) was thus not so much the advantage of easier access to information in my own herbarium, but more a personal interest in the task itself. Nevertheless it provides a prototype on which larger systems could be based, and gives an idea of the sort of tasks for which a micro can be used.

Building a herbarium maintenance system on a microcomputer is not a project to be tackled lightly. It may take a fair amount of time and money to get a working system off the ground (the first BACCHUS system was born about 8 years ago, and is still being added to), so partly it depends on how important to you the flexi-

lity and reliability of the system is, and how much you are prepared to spend. We tend not to hear about the failures, and most large organisations that have tackled this problem are using very powerful microcomputers or mainframes.

Using a modern database package, such as dBaseIII+, is certainly the quickest way of doing it, but this may not provide sufficient flexibility to do what you want, and the report writing features of such packages are often not very good. Writing the system in a conventional programming language, such as BASIC, will take a long time, especially if you are inexperienced in such matters, and means creating a set of inter-dependent files where a change to one may mean an associated change is needed to others - which adds to the complexity of the programming, and thus the likely programming errors.

We are in the fortunate position today of being able to benefit from the efforts of those who have gone before, and there are several packages available in the IAB Software Library. Some

Of these are built round the dBase proprietary software package, which costs several hundred pounds, but if you have access to this, then it is a useful starting point. Hand-coded products aren't necessarily any better, but may be more flexible in what they can offer. (With dBaseIV, it is possible to distribute programs together with a free 'runtime' version of dBase, which means that you don't need to own a copy of dBase to be able to use dBase programs.)

Before starting the development of any computer system it is important to know just what it is you are trying to develop. Even when you are developing something for yourself, it is worth spending the time to produce a statement of your requirements before you begin. It is also worth analysing the data you are dealing with. Many problems of both development and maintenance of computer software stem from inadequate preparation, and it is generally recognised that less than 20% of software development time should be spent writing programs - the rest is spent in the preliminary stages of deciding whether the project is worth doing and producing a statement of requirements, and also in the final acceptance test of the system. There is also a high cost of failure in software products - manpower is expensive, and so must be used efficiently and effectively. The cost of failure will be significantly reduced the earlier in the development a problem is found: getting the specification right at the outset may save weeks of time going down the wrong route, which may only be discovered when the system is complete.

For all these reasons, it is worth looking at the environment in which herbarium management systems exist. If we look first at bryological functions in an institution, they might include herbarium maintenance, taxonomy, floristics, ecology, identification, bibliography, report production and administration - though it isn't easy to see where any overlaps exist, and report writing for instance might be supported centrally. Each function would need to be split down further before any overlaps could be recognised. The herbarium maintenance function for instance could split into loans, accessions, and so on, but it may not be clear where the identification function belongs - and it may also not be clear how an event in one area is signalled to another. It ought to be easier to look instead at the information in the organisation, because this isn't so much affected by preconceptions

about organisational functions.

Concentrating on the herbarium data, we might be interested in the following entities ('entity' is data analysis parlance for 'something we are interested in, and want to store information about'):

<u>Entity</u>	<u>Possible attributes</u>
Author	
Authority	
Bibl. details	
Borrower	name, address, loan nos.
Collection	date, habitat, coll et det, no.
Description	text, diagrams
Herbarium	name
Label	
Loan	date, to whom, no. of specimens, historical record
Site	grid ref, lat/long, alt, location
Synonyms	
Taxon	name, catalogue no.
Taxonomy	

These could possibly be related as shown in the figure on p. 9.

Even if your interest is not as wide as this, it is important to be able to exchange information with other botanists, and being aware of the logical structure of herbarium data may make this task very much easier.

The IAB Software Library contains a number of programs of interest to herbarium managers:

#3	Herbarium label program
#37	Labels3
#38	Herb
#53	Determination Lists
#64-65	Herbarium curatorial prgms.
#103	Loan
#104-5	Herbarium manager
#118-9	BACCHUS
#154-5	Herbarium label and record book program (continues #3)

This issue of the Bryological Times is a contribution of the Computer Techniques Column.

Column editors :

**Dr. J.-P. Frahm and
B.J.O'Shea**

Distribution mapping using DMAP

by B.J. O'Shea

The DMAP system was developed by Dr. Alan Morton to help with a project to map the flowering plants of the Welsh county of Montgomery, but the examples included here show data provided by Dr. Jeff Bates, from the bryophyte flora of the English county of Berkshire, of which Reading is the main town. DMAP was written using Borland's Turbo Pascal 4, using an Amstrad PC1512.

The factors that please me about DMAP are two-fold: firstly, it is a simple tool, fit for a specific purpose, with no irrelevant 'bells and whistles', yet with the ability to produce very high quality maps on the cheapest printers; secondly, it is a tool of great flexibility, but still very simple to use. This latter point is particularly important, as its facilities are generalised to allow it to be used on any gridded map, from country scale down to local counties, often in the UK mapped to a 4 square kilometre grid, or sometimes smaller. Its value and utility is thus international, provided maps have an orthogonal grid.

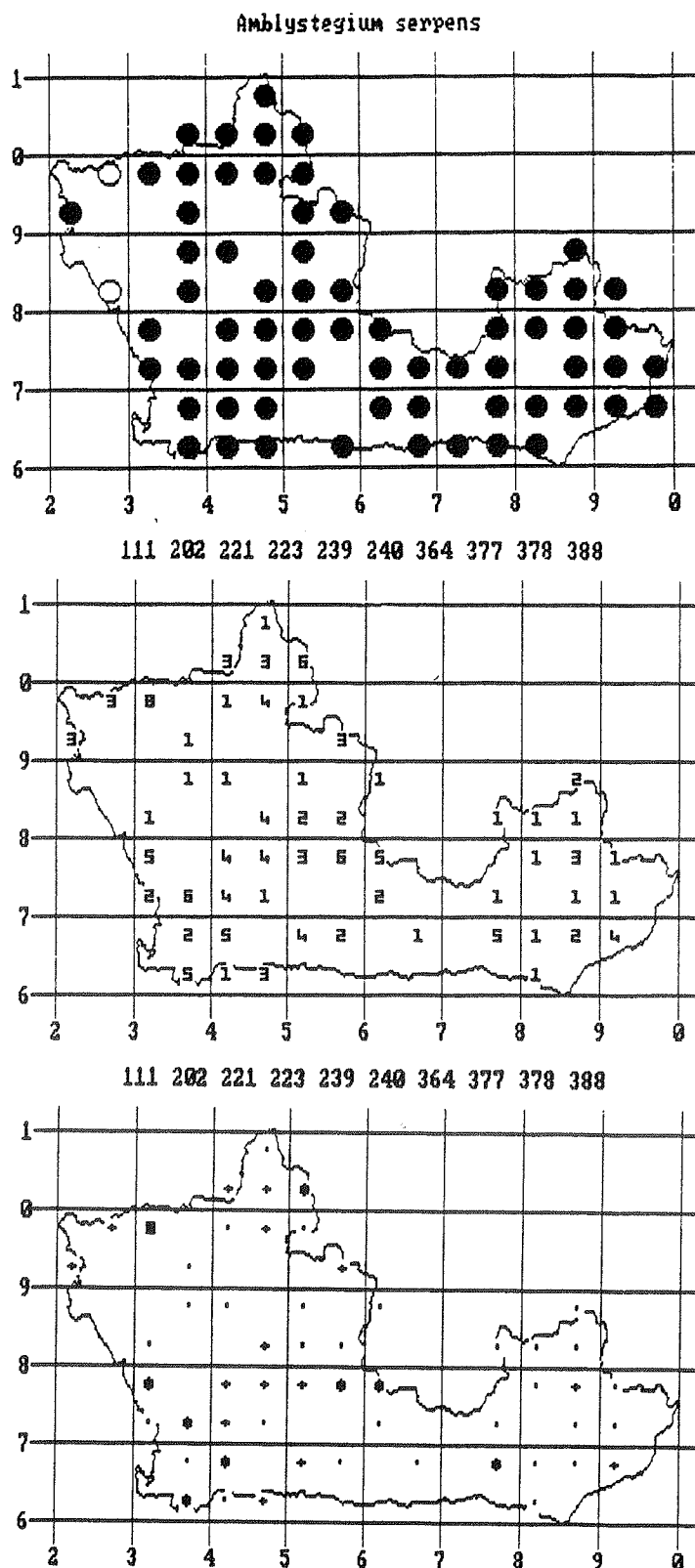
The program has been developed to run on IBM-compatible microcomputers, and dynamically interrogates the graphics hardware to determine its type, and will run on CGA, EGA, VGA, Hercules and other displays. As well as being displayed on the screen, the maps can be printed by a dot matrix printer (800 x 400 pixels) and certain types of inkjet and laser printers, (using up to 800 x 800 pixels). The program handles all the scaling from the differing screen and printer display types.

Data files must be supplied by the user to provide: parameters controlling the size and scale of the map; a boundary outline held as grid references; the names of species (or other plotted entities, e.g. habitats); and the distribution data, supplied as grid references. Full descriptions of these files are supplied in the documentation that comes with the program.

The function of the program is for distribution or coincidence map plotting; it does not pretend to anything else, and does not for instance pro

Figures

1. Any map showing grid
2. Coincidence map using symbols
3. Coincidence map using numbers



Vide a database, although there is a modest edit program that will output DMAP format files. The data structure however is very simple and would be extremely easy to create from other databases. Ideally DMAP would be used as an adjunct to the main floristic database, and could for instance be called using parameters indicating the data files to be used.

Other features of the program are as follows:

- a grid is automatically superimposed on a map, numbered on the south and west margins (although a map can be displayed without a grid if required) (see Figure 1)
- the extent of the gridded area can be user-defined
- the scale of the plotted map can be user defined
- the map is automatically centred on the screen
- the size of the distribution map plotting symbol (which can be round or rectangular) can be user-defined
- a range of symbol types, sizes and fillings can be selected for individual records in the distribution data file (see Figure 2)
- the coincidence mapping facility can be used for plotting the quantitative distribution of groups of species for 'habitat' mapping, plotted using either symbols (Figure 3) or numbers (Figure 4)
- grid reference data can be resampled for scales greater than that supplied on the data file: for example, if 1km grid references are supplied, distribution maps for 10km could be produced
- printing can be done on 9-pin or 24-pin matrix printers, and inkjet or laser printers that support the Hewlett-Packard Printer Command Language (PCL). These printers are capable of producing publication-quality graphs.

The author of DMAP would be pleased to hear from anyone interested in using the program, and as well as providing a copy of the program and documentation, will also provide whatever assistance he can over the phone. In order to defray the costs of developing DMAP, and to support its continued development, a contribution of £10 is requested. A copy of the program is distributed free of charge via the IAB Software Library, but this is not the latest, and is not supported by Alan Morton. Please contact Dr Morton at: Department of Pure and Applied Biology, Imperial College, Silwood Park, Ascot, Berks SL5 7PY, UK.

The Proceedings of the IAB Computer Workshop will be continued with more contributions in the next issues of the Bryological Times.

TROPICAL BRYOLOGY, the first bryological journal on diskette

The journal TROPICAL BRYOLOGY is also available as an ASCII file on diskette in any MS-DOS format. The computerized version of the journal has several advantages over the printed version:

- It allows the reader to search for keywords, authors, references etc. with the DOS command 'find' or (more easily) with special programs. FSRCH (IAB software library # 218) can search for any string or combinations (and/or) of keywords.
- It allows the readers also to copy parts of the text (descriptions, lists of specimens, references) for their own files instead of retyping them.
- References can be used for a personal database, especially in combination with unstructured database programs (such as Infobase, IAB software library # 6, or Instant Recall, IAB software library #71). These programs recognize each reference as a record and allow you to search, select and print references.

Tropical Bryology on diskette is available for subscription price.

SOFTWARE LIBRARY Program of the month: #136 Expressgraph

