



## CONTENTS

|  |    |
|--|----|
| <b>EDITORIAL</b> .....   | 1  |
| <b>TURNING TAXONOMIC DATA INTO INFORMATION AT HERBARIUM BOGORIENSE</b><br>by J. Jarvie .....                         | 2  |
| <b>A DELTA DATABASE PROJECT FOR THE PLANT GENERA OF WESTERN AUSTRALIA</b><br>by T.D. Macfarlane .....                | 4  |
| <b>DELTA AND HYPER-TEXT MARKUP LANGUAGE</b><br>by A.R. Chapman, N.S. Lander, T.D. Macfarlane and M.J. Dallwitz ..... | 5  |
| <b>PRACTICAL DELTA PROBLEMS AND SOLUTIONS:</b>   |    |
| Discussion from the DELTA software listserver .....  | 8  |
| <b>Suitability of DELTA for fruit/seed key. Use on Macs</b> .....  | 8  |
| <b>Troubleshooting DELTA programs. Sending DELTA files by email</b> .....  | 9  |
| <b>Advantages of using a separate subdirectory for each data set</b> .....   | 10 |
| <b>Merging DELTA data sets</b> .....   | 10 |
| <b>Combining DELTA items</b> .....   | 11 |
| <b>INTKEY packages for lichens. Using the same characters at different taxonomic levels</b>                          | 12 |
| <b>Recording host ranges, etc., in DELTA</b> .....   | 14 |
| <b>Regional subsets of taxa in INTKEY</b> .....  | 14 |
| <b>INTKEY: direct access to characters which are illustrated</b> .....   | 16 |
| <b>Hiding 'inapplicable' character states</b> .....  | 17 |
| <b>Using DELTA to produce non English descriptions and keys</b> .....  | 17 |
| <b>Using INTKEY over the Internet</b> .....  | 18 |
| <b>DELTA and Pentium</b> .....   | 18 |
| <b>Screen capture</b> .....  | 18 |
| <b>RECENT PUBLICATIONS</b> .....   | 19 |
| <b>RELEASE OF INTKEY FOR WINDOWS</b> by T.D. Macfarlane .....  | 20 |

## EDITORIAL

Various circumstances prevented the preparation of last October's issue, but here we are again. New printing and distribution arrangements have been made, which relieve financial pressure on the Newsletter. We thank Dr Ebbe Nielsen and the CSIRO Division of Entomology for these arrangements, and we hope the associated material will also be of interest to you.

The bulk of this issue consists of a reproduction of discussion from the temporary DELTA-L Internet listserver. Questions of the Frequently Asked kind are answered, and there are a number of detailed treatments of practical problems, including worked examples. We believe that the dissemination of this type of information, where various DELTA format and software features are discussed in detail in the context of real data, is an important role for the Newsletter. We also felt that the DELTA-L discussion merited a wider exposure and a place in the growing body of similar information which is building up in this Newsletter. In the future we plan to index the Newsletter's contents so as to provide a ready means of locating articles that provide help on particular aspects.

Other features of this issue include descriptions of two DELTA projects on flora; discussion of a soon-to-be implemented HTML translation facility in CONFOR which will allow automatic production of hypertext descriptions, an area of growing importance; and notices of relevant recent references. The regular columns Personal Profile and Spotlight on DELTA Features have been held over until next issue. Contributed articles, letters, notices of meetings, descriptions of projects and so forth are always welcome, so plan now to get yours to us by 1 October next.

Terry D. Macfarlane\*, Nicholas S. Lander, Alex R. Chapman  
Western Australian Herbarium  
Department of Conservation and Land Management  
PO Box 104, Como, WA 6152  
AUSTRALIA  
Telephone +61 9 3340500 Fax +61 9 3340515 Email (via M.J. Dallwitz: INTERNET md@ento.csiro.au)

\* Based at Manjimup Research Centre, Department of Conservation and Land Management, Brain Street, Manjimup, WA 6258

## Turning Taxonomic Data into Information at Herbarium Bogoriense

Jim Jarvie\*

The Arnold Arboretum of Harvard University, 22 Divinity Avenue, Cambridge, MA 02138 USA

\* Current mailing address: Herbarium Bogoriense, Jalan Juanda 22, Bogor - 16122, Indonesia  
E-mail: 0006486174@mcimail.com

Indonesia has a rich and diverse biota, but taxonomic information about it is sparse. Fundamental data sources, including revisions and monographs, tend to be locked in English with a variety of formats, which obstructs their synthesis. Furthermore, many are already out of date. This lack of data, and the difficulties data users have in translating available data into information understandable by their clients, is a bottleneck hindering applied uses of taxonomic data to areas such as conservation. Aside from obvious species such as orang-utan, rhinoceros, tiger and some birds, taxonomic data are generally not available or ignored by conservation planners for the majority. It is ironic that whereas a large part of the global focus on conservation has resulted from the advertisement of biodiversity destruction, very little detailed biodiversity information is available or ever used for conservation efforts concerning the majority of the species around us.

Various projects are addressing the disconnection between taxonomic data and their critical use in conservation, often in the form of databases and their applications. Among these is a collaborative project between the Arnold Arboretum and Herbarium Bogoriense, funded by the National Science Foundation. Its basic objective is to develop a woody plant flora for the Bukit Baka-Bukit Raya National Park in Kalimantan. The wider goal has become to develop an information base translating taxonomic data into useful and used information. Whereas its geographic focus is small, we hope it will act as a pilot to stimulate comparative studies in other areas in a compatible way. The project has highlighted elsewhere the problems with current 'static' floras and possible ways around these (Jarvie & Welzen 1994). Presented here is a brief description of our approach.

In the first instance the project required a database to hold field data from our collections and those pertaining to the trees and shrubs of West Kalimantan from the herbarium. It quickly became apparent that given our location the system had to be small and "home- serviceable". A simple database system based on Paradox 4 with English and Indonesian input forms was devised. This is linked to a WP6.0 for Windows merge facility for label production. The system currently contains data from over 3,000 herbarium specimens plus another 3,500 from the project's own collections in and around the park. This system is a holdover until the development of a larger system encompassing GIS facilities with rapid access to expert assistance.

The project has discussed outputs with a variety of end- users. These included collaborating researchers; government conservation bodies; NGOs; university staff and lectures. It became quickly apparent that we had a range of users with a variety of backgrounds. Two outputs were highlighted. A printed, but flexible, flora and a reliable method to identify the taxa in the area.

### Use of DELTA

DELTA has been extensively used in this project. We have focused at the generic level. DELTA character lists for the species are for later workers! The keystones of our system are two 244 character lists, one in English and one in Indonesian (Jarvie et al. 1994). This is adapted from the Watson and Dallwitz (1991) character list for the families of flowering plants. Characters pertaining to features which were not going to be used, such as chemical components, were removed. Characters important in forest identification, such as slash and buttress features, were added.

### Creating a Flora

One use of DELTA is for the flora production. Data is input through a Paradox 4 entry system with a macro which outputs data as an ITEMS file. The directory then outputs English or Indonesian family and generic descriptions. These will form the basis for a modular "spring-leaf" format flora augmented by written diagnostic keys and scanned drawings or vegetative features of importance. We have identified the need for different sorts of outputs for genera. One is for descriptions with a wide scope for university use. The other is for restricted outputs of vegetative characters which field staff can take with them to approximate identifications of certain trees.

### Creating an Interactive Key

Under development is an interactive key to the woody genera of the park. It uses 113 characters from the original "stable" of 244. There are currently 202 genera included and about 150 of these are illustrated. A prototype version is available (please send requests to the Bogor address or even better, the e-mail address), but users should be aware that there are still errors. We are currently improving this version in the light of the constructive criticism we have received. Coding errors have been found and we need to tighten up the resolution of our images. This is the first interactive key to the woody genera of the magnificent hill Dipterocarp forests of Borneo. We hope it can later be expanded to the swamp forests and to the mangroves.

Systems assisting identification of the biodiversity of S.E. Asia need to be developed and advertised. They are needed, but this need is often unnoticed until workable products are developed and passed in the context of the immediate issues facing the loss of the biodiversity they describe. Hopefully these systems will be improved and expanded, and at the same time encourage the wider community to develop more efficient conservation strategies based on an appreciation of the taxa they are dealing with.

### References

- Jarvie, J.K. and P. v. Welzen. 1994. What are tropical floras for in S.E. Asia? *Taxon* 43:444-448.
- Jarvie, J., U.W. Mahyar and Ermayanti. 1994. A DELTA character list for family and generic descriptions of plants in English and Indonesian is now available. *Tropical Biodiversity* 2:280-281.
- Watson, L. and M.J. Dallwitz. 1991. Automated descriptions of the families of flowering plants. *Australian Systematic Botany* 4:681-695.



## A DELTA Database Project for the Plant Genera of Western Australia

T.D. Macfarlane

A DELTA database project for the Angiosperm genera of Western Australia is in progress for the Department of Conservation and Land Management (CALM), centred at the Western Australian Herbarium, Perth. This is a brief note to announce the existence of this project, which is managed by T.D. Macfarlane, N.G. Marchant and L. Watson. Some other people who have worked on aspects of the project include H.R. White, T.R. Lally, J. Gathe and J.R. Wheeler.

The Western Australian flora is very diverse, having about 12, 000 species with a high level of endemism at the species and genera levels. The complement of ca 1300 genera, including naturalised plants, represents both tropical and temperate floras, and is a little more than half the total number of genera in Australia as a whole. Western Australia is not fully covered by modern Floras, a situation which will not be remedied for many years. Even those now available are subject to obsolescence, are difficult to revise, and expensive to publish. Yet there is a high demand for current, authoritative information from government agencies concerned with land management, biological conservation and environmental protection, and from other sections of society. A database providing floristic information and a means of identification is a solution to the problem of information currency. And there will be improved resource efficiency over the long term by the use of a multiple-use data format, DELTA, through overcoming repetitious data gathering and entry.

With Angiosperm families catered for by the Watson & Dallwitz (1991, 1994) database, the logical next step is to cover the genera, for which a Flora was in any case a long-standing plan. Eventually it is planned to produce a database covering the whole flora to species level, but this will be a different and hopefully more widely based cooperative venture. Meanwhile various databases are underway at the species level as parts of other projects.

Our character list is largely taken from the Angiosperm Families database of Watson & Dallwitz, and in fact we have incorporated data for families into the database, so that there are data present at family level even where the genera of a given family have not yet been recorded. Because we wanted to record maximum useful data and provide the ability to identify specimens to genus, we have extended the standard character list with such further characters as are needed in particular families in order to differentiate the genera. These additional characters are incorporated into a logical position in the character list soon after they are added, but they are recorded only for those genera for which they are relevant.

To date we have entered data for over 25% of the genera which we plan to record. We are excluding families for which published DELTA databases of genera already exist, e.g. Poaceae. It is planned to complete initial data entry in 1997.

As a consequence of using the DELTA data format, the output from the database can take a variety of forms. We plan to provide INTKEY databases for information retrieval and identification, printed descriptions tailored to the required detail and scope, and hypertext descriptions made available on our internal network, and ultimately, on World Wide Web. Images are to be included, and some are already in.

### References

- Watson, L. & Dallwitz, M.J. (1991). The families of Angiosperms: automated descriptions, with interactive identification and information retrieval. *Australian Systematic Botany* 4: 681-685.
- Watson, L. & Dallwitz, M.J. (1994). *The Families of Flowering Plants. Interactive Identification and Information Retrieval*. Booklet and CD-ROM for MS-DOS. (CSIRO Publications: East Melbourne.)

## DELTA and Hyper-Text Markup Language

A.R. Chapman, N.S. Lander, T.D. Macfarlane and M.J. Dallwitz

While professional taxonomists appreciate the flexibility and functionality of the DELTA system and products such as CONFOR and INTKEY, there is a necessary complexity inherent in these tools which can obscure the information content of the data therein. With the advent of the World Wide Web, a graphical hypertext environment for documents on the Internet, a simple yet extensible interface became available for use within and between institutions and individuals throughout the world.

The vision is to be able to output DELTA data in a format readable by any of the World Wide Web browsers (such as *Mosaic* or *Netscape*) and containing hypertext links from each item name to a natural language description or taxon image. This would create a nest of inter-related HTML documents on a particular taxonomic group which was navigable by a naive user to easily retrieve taxonomic information and potentially available worldwide.

Formatting documents for the World Wide Web involves the embedding of Hyper-Text Markup Language or HTML codes into the document text to indicate heading levels, list types and titles, links to other text, image or sound files, and character formatting. HTML is a subset of the ISO's Standard Generalised Markup Language (SGML) which provides a software-independent means of encoding formatting in a document.

In effect this requires the definition of a new directives file which would be similar to, and contain many of the same directives as, TONAT and TONATT. However, a number of new directives specific to the functionality of HTML documents are needed. What follows is a description of some of the requirements we think are essential for preparing HTML output from DELTA coded data. Note that sufficient of these requirements are met by the impending release of the DELTA suite of programs to allow the output of useful HTML descriptions from DELTA-coded data.

### Text

The current CONFOR directives ITEMS SUBHEADINGS and TYPESETTING MARKS allow the incorporation of additional text and typesetting marks in various contexts in natural-language descriptions. The same marks are placed in all instances of a given context. It is a simple matter to change the typesetting marks from the TYPSET codes used in the current sample files to HTML codes.

Several more 'contexts' have been added to the TYPESETTING MARKS directive in the next program release. Those particularly relevant to HTML are 'after each taxon description' (supplementing the existing 'before each taxon description'), 'at the start of each output file', and 'at the end of each output file'. This allows, for example, the placement of an icon before each description, a horizontal rule after each description, and headers and footers at the start and end of each file. The headers and footers could contain introductory text and links to an index page and/or a home page. Ideally, the header should contain a title to be displayed in the Title Bar of the Web browser window, but the nature of the title is currently limited by the requirement that the same marks must always be used (in a given CONFOR run).

Angle brackets are used by HTML to encapsulate its commands and could conflict with the use of angle brackets in the DELTA format. This kind of problem was anticipated when the TYPESETTING MARKS directive was originally implemented, and an alternative delimiter may be defined, allowing the use of any DELTA delimiters within the typesetting marks. The planned new CONFOR will have general mechanisms for delimiter substitution and different typesetting systems, but for the time being, any typesetting marks to be embedded in the main data files (characters and items) will have to be in TYPSET format, and substitutions will be made by CONFOR during output of the natural-language descriptions, or by editor macros afterwards. Prototype systems which have been developed for HTML and RTF (Microsoft Rich Text Format) use the latter method.

Ultimately, a simple means of controlling the formatting of the output text in terms of HTML Document Type Definitions needs to be found. And problems with the representation of foreign characters and symbols will need to be overcome.

## Images

The currently available \*CHARACTER FOR TAXON IMAGES directive can be used to code the name of the image file(s) to be associated with the description of a particular taxon. The DELTA development team are currently reformulating directives regarding taxon images. Thus an alternative format for this data will shortly be available which identifies the name of a taxon images file (analogous to the current character images file). These two mechanisms are equivalent but mutually exclusive.

Note that the point of insertion of the image (or images) may be controlled simply by the placement of the taxon image character in the natural language output. Captions for image files can be recorded in the same way as those for use with the forthcoming INTKEY for Windows package, namely by the use of INTIMATE, a new DELTA utility for editing the overlays to be associated with a particular image file.

In the WWW environment it will be very useful to signify the availability of an image for a taxon without committing the end-user to the inevitably slow download of an HTML document with a large GIF image embedded. Thus allowance must be made for inserting a hypertext link to a GIF image which will be loaded on following that link. In the forthcoming DELTA release the default link is the image file name.

Ultimately it would be desirable to be able to declare optional variables to control the type of link object, for example an alternative text string to the image file name, or the number of another taxon image character containing a reference to a *thumbnail* GIF image. This latter option would allow the user to see a small version of the image which when selected would download the full image for perusal (see Figure 1).

## Files

With large data sets containing many descriptions it will be desirable to split the HTML output into a number of files, the default being that the description for each item is output to a separate file. This is possible by means of the currently available directive \*NEW FILES AT ITEMS. A simple sequential numbering system is currently used to name each resulting file (e.g. *tonat001.htm*, *tonat002.htm*, ...). If this directive is not used then the output all goes to a single file.

Ideally (and in either case) the main output file would need to include a list of item names (similar to the output from PRINTN) which contained hypertext links either locally to the appropriate item description in the file, or to each of the files produced using the \*NEW FILES AT ITEMS directive.

Ultimately, it would be desirable to provide a mechanism whereby output files for each taxon are uniquely named, perhaps according to a system of taxon codes for use in corporate information systems. This could be readily achieved by declaring a character for the taxon code upon which the output file names are to be based.

## Other Considerations

Another area of DELTA where the adoption of HTML formatting might usefully be explored is in the production of polychotomous keys using the KEY program. By using hypertext links from each lead in a key a user could navigate to the next lead, to a character definition or image to aid in the use of the key, or from a resulting taxon name to its description or image.

For taxonomists wishing to distribute their DELTA datasets and numerous images to a wider audience, then CD-ROM is almost certainly the medium of choice. We see much scope for bundling and presenting the data not only with the latest version of INTKEY for Windows, but with a well-designed set of hypertextual natural language descriptions

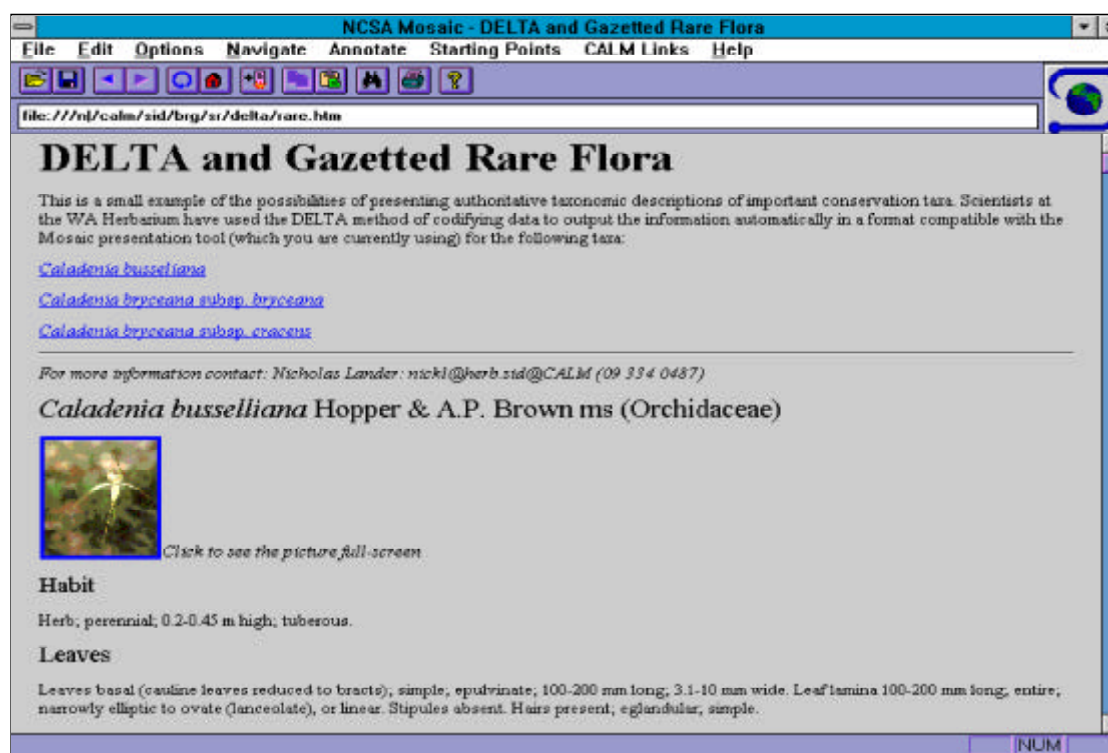


Figure 1. An example HTML document, including a related image, which demonstrates some of the final output we envisage from a TOHTML directives file.

navigable using a freeware Web browser, such as NCSA's *Mosaic*. Not the least advantage of this is that the default image format acceptable to both INTKEY and HTML is CompuServe's Graphics Image File (GIF) format, hence duplication of image files, which are usually quite large, is avoided. Conceivably, owners of the CD-ROMs could keep these datasets current by down-loading the latest DELTA files from the Internet.

As more taxonomists gain full access to the Internet then, by using the methods we have outlined here, the hypertextual descriptions generated from their datasets could be readily presented live on the Internet and kept up to date for use by the World Wide Web user community.

To illustrate this concept an HTML version of Watson & Dallwitz' Angiosperm family descriptions, automatically generated from DELTA codings, is available at these URL's: <http://muse.bio.cornell.edu/delta/angio/extra/index.htm> and <http://spider.ento.csiro.au/delta/angio/extra/index.htm>. It may integrate well with the *Tree of Life* project being implemented by David R. and Wayne P. Maddison (University of Arizona) on the Web, in which they are endeavouring to focus contributions from around the world to create a central map of the phylogeny of organisms, from kingdom to species. This presentation currently makes use of an earlier manually marked up version of the Angiosperm family descriptions. The URL for their Web site is: <http://phylogeny.arizona.edu/tree/phylogeny.html>.





## Practical DELTA problems and solutions: discussion from the DELTA Software Listserver

The following composite article is a selection from the DELTA Software Listserver on the Internet, the origins of which are explained in the first message. The text has undergone minimal editing, mainly for deletion of superfluous header material, corrections of obvious errors, and cross referencing footnotes. The editors thank Mike Dallwitz for providing the raw material, and Phillip Hoover and Derek Ellis for providing the forum which produced this useful body of information. The text is reproduced here because it is considered to be of wide interest, and Dallwitz's replies were written with generality in mind.

Sender: Biological Systematics Discussion List  
<TAXACOM@cmsa.Berkeley.EDU>  
From: Phillip Hoover <phoover@SOL.UVIC.CA>, 12  
Dec 1994  
Subject: DELTA software mailing list  
To: Multiple recipients of list TAXACOM  
<TAXACOM@cmsa.Berkeley.EDU>  
DELTA Software Listserver

### DELTA Software Mailing List

A trial DELTA (DEscription Language for TAXonomy) users e-mail discussion list is being established in December of 1994 using the listserver at the University of Victoria (British Columbia, Canada). The list will be of a temporary nature, running until April 1994.

The purpose of this list is to allow discussion regarding the use of the DELTA system for its users and those interested in learning the use of the system. It could allow posting announcements of new upgrades and additional programs for use with the DELTA system and meetings and workshops of interest to DELTA users. We are setting it up for the education of our students in a final year undergraduate course: Taxonomy and Biodiversity.

There is no cost for subscription.

Messages sent to the list DELTA-L@UVVM.UVIC.CA will be forwarded to all members subscribing to the list. To subscribe, send a message to the listserver LISTSERV@UVVM.UVIC.CA (or LISTSERV@UVVM.BITNET) with a message in the text saying:

subscribe DELTA-L your name

The subject line should be left blank. Questions regarding the list can be sent to the list manager PhillipHoover (phoover@sol.uvic.ca).

From: Phillip Hoover <PHOOVER@SOL.BITNET>

Here is the first posting to this list. The questions are fairly basic and I think I will start to put together a FAQ (Frequently Asked Questions) file that can be sent to new and interested users of the DELTA system.

I do think that DELTA would be useful for Susan Mazer's project. Would any else like to reply to the other questions?

————— Forwarded message —————

**Subject: Suitability of DELTA for fruit/seed key. Use on Macs**

From: Susan Mazer  
(E-Mail mazer@lifesci.lscf.ucsb.edu), Department of Biology, University of California, Santa Barbara, CA

I am currently working on a photographic field guide to the fruits and seeds of about 1500 Peruvian lowland rainforest species, and I'm interested in the possibility of using Delta to produce a dichotomous key to allow the field identification of these taxa.

A few elementary questions (I clearly don't know much about Delta):

Is this an appropriate/possible application of Delta?

Is Delta now available for use on Macintosh computers, or only for DOS?

Where can I look into the purchase of the program?

————— End forwarded message —————

From: Mike Dallwitz <miked@ento.csiro.au>, 13  
December 1994



This is an appropriate application of DELTA, though I strongly suggest that you develop an INTKEY (interactive identification) package instead of (or at least in addition to) conventional keys. INTKEY could comfortably handle 1500 species in a single 'key'.

Our DELTA programs are only available for MS-DOS. A MS-Windows version of INTKEY is almost ready for release. The programs will run on DOS and Windows emulators on a Mac<sup>1</sup>. We recently did some tests on a Power Mac 7100/66AV with 16MB of RAM. The speed in DOS/Windows emulation is roughly 15% that of a 50MHz 486. The substantially lower speed is no problem for most INTKEY operations, which complete almost instantaneously. However, the time taken for operations which involve scanning the whole data matrix might be regarded as inconvenient. On John Lawrence's beetle-larvae data (180 characters, 385 taxa), a check of the main DELTA data files took 167sec, and the INTKEY BEST command took 145sec in DOS and 120sec in Windows. (Bear in mind, however, that you don't need to wait for BEST to complete before choosing a character.)

The DELTA programs, and several data sets, are available by gopher or anonymous ftp from the following Internet hosts:

muse.bio.cornell.edu (directory: /pub/delta)  
spider.ento.csiro.au (directory: /delta)

The file Index.txt (note the upper-case I) contains a list of the available programs and data. To view this file during an ftp session, give the ftp commands 'get Index.txt' and '! more < Index.txt'. Most of the subdirectories of delta contain a file Download.1st which contains information about downloading and installing the files in that subdirectory. Files other than .1st, .txt, and .ps must be transmitted in binary mode (set by the ftp command 'binary').

The programs are supplied with documentation files, sample data, and a list of references. The conditions of use are in a file DELTA.USE, and the prices in DELTA.REG. These files are within the self-extracting archive file DELTA1@.EXE.

Updates are announced on the Internet via the Taxacom mailing list. To subscribe to Taxacom, send the message

subscribe taxacom your-first-name your-last-name  
by email to listserv@cmsa.berkeley.edu.

Mike Dallwitz Internet md@ento.csiro.au  
CSIRO Division of Entomology Fax +61 6 246 4000  
GPO Box 1700, Canberra ACT 2601, Australia

**Subject: Troubleshooting DELTA programs.  
Sending DELTA files by email**

From: Mike Dallwitz, 19 Dec 1994

If you have a problem running any of the DELTA programs with your data, the first step is to try to solve it by referring to the User's Guide, particularly Section 3.3 (error messages), and the descriptions of the directives that seem to be giving trouble (e.g. Sections 2.1, 2.2, 3.6). Comparing your files with the sample files supplied with the programs can often help to pinpoint the area of difficulty. If none of this helps, send us email describing the problem. If there was an error message, append the 'listing' file (e.g. TONAT.LST) to your mail.

If we can't solve the problem from the information you send us, we will ask you to send your files, so that we can try to reproduce the problem here. You should send the DELTA-format files and directives, SPECS, CHARS, ITEMS, TOKEY, TONAT, TOINT, etc., plus INTKEY.INI. Don't send binary files like ICHARS, IITEMS, etc. - we can generate them from the DELTA files.

You can append the DELTA files to an email message, but before doing so, please compress the files using LHA or PKZIP, and then UUENCODE the resulting .LZH or .ZIP archive file. The archive file will be smaller than the original files, and, more importantly from our point of view, we can regenerate the original files without manual editing. UUENCODE converts the binary archive file to a text file, which is necessary because the mail system can't handle binary files. LHA, PKZIP, and UUENCODE are available on the DELTA gopher/ftp sites:

muse.bio.cornell.edu (directory: /pub/delta/util)  
spider.ento.csiro.au (directory: /delta/util)

Here is the table of contents of the util subdirectory.

|               |   |
|---------------|---|
| lha213@.exe   | Free file-compression program for DOS.                                      |
| pkz204g.exe   | Shareware file-compression program for MS-DOS (PKZIP).                      |
| selfex@.exe   | Batch files for generating self-extracting files, using LHA.                |
| uudecod@.exe  | uudecode & uuencode for DOS (complete)                                      |
| uudecode.1st} | Text files for generating uudecode and uuencode for MS-DOS. Use these files |
| uudecode.bas} | only if you cannot download the binary file                                 |
| uudecode.doc} | uudecode@.exe, or if you want to send the programs by email.                |
| uuencode.uue} | See uudecode.1st for instructions.  |
| winedit@.exe  | Shareware text editor for MS-Windows.                                       |

Mike Dallwitz Internet md@ento.csiro.au

**Subject: Advantages of using a separate subdirectory for each data set**

From: Gerhard Rambold  
(rambold@botanik.biologie.uni-muenchen.de)

"The reason why I use POA\_I.INT instead of ICHARS is because there are various INTKEY keys which have to be located within one and the same subdirectory, to avoid multiple installation of INTKEY.EXE etc. I start a key with, e.g., INTKEY POA.INI. POA.INI has POA\_C.INT and POA\_I.INT as data files, and so on."

From: Mike Dallwitz, 15 Dec 1994

The DELTA convention of keeping data sets in different subdirectories is convenient, because it minimizes editing of directives files (because the file names can be the same). It is possible to arrange things so that this is also convenient for using INTKEY. You need only one copy of INTKEY.EXE, as long as its directory is on the path. If you have each genus in a subdirectory, you can set things up so that you can enter TO POA, then INTKEY. For example, if your genus directories are subdirectories of \GRASSES, you should put

```
SET CDPATH=.;.;\GRASSES
in AUTOEXEC.BAT. To make starting INTKEY even easier, you could set up batch files in a directory on the path, e.g. POA.BAT -
```

```
@ECHO OFF
CD \GRASSES\POA
INTKEY
```

so that you could just enter POA at any point.

The Windows version of INTKEY selects data sets via an index, which can contain descriptive entries, e.g. Species of *Poa*<sup>2</sup>.

From: Phillip Hoover <phoover@SOL.UVIC.CA> 15 Dec 1994

When teaching students how to use DELTA we must keep each student working on their own files in separate subdirectories as there are a number of DELTA users on each computer.

**Subject: Merging DELTA data sets**

From: Phillip Hoover <phoover@SOL.UVIC.CA>

A student has produced two different DELTA data files sets for two different genera but now wishes to combine them to use with Intkey or for the production

of printed keys. What is the most effective way to do this? In the past I would get them to create a another subdirectory and enter new data files. Normally students are only dealing with 6-10 taxa and a dozen characters so it is a fairly easy job.

From: Mike Dallwitz 16 Dec 1994

The new DELTA system will incorporate a simple method of merging data sets. At present, it can only be done by the following procedure. Note that there is little point in combining data sets unless they share at least some characters.

1. Reorder both character lists, and the items etc., so that the shared characters are at the starts of the lists.

2. To the end of the second list, add dummy characters equal in number to the non-shared characters in the other list. It is easiest to make these UM characters, of the form  
#n. / 1. / 2. /

It may be possible to make your text editor generate the required sequence of numbers automatically (TED, supplied with the DELTA programs, can do this via the T command). Don't forget to increase the NUMBER OF CHARACTERS in the SPECS.

3. Reorder the second list, and the corresponding items etc., so that the dummy characters are immediately after the shared characters; i.e., so that they are aligned with the non-shared characters of the first list.

4. From the second list, delete the shared and dummy characters. Append the remaining characters to the first list. (Check that the sequence of character numbers in the combined list is correct.)

5. Form a combined SPECS from the SPECS of the first list, and the reordered SPECS.NEW of the second.

6. Form a combined ITEMS from the items of the first list, and the reordered ITEMS.NEW of the second. If the two sets of items do not describe the same entities (e.g. each set of items describes the species of a different genus), the second items file (less the directives at its start) can simply be appended to the first. If the two sets of items describe the same entities, each item from the second list must be inserted (less its item name) after the corresponding item in the first. (A TED box program, ADDLIST, is available for this purpose.)

7. TIDY the combined items. If the two sets of items described the same entities, it will be necessary to use an ACCEPT DUPLICATE VALUES directive (presumably the shared characters were coded the same in both sets of items).

#### Example

The comments are included to identify the data set from which the information came. Assume that the character lists have already been ordered (step 1) so that the shared characters are at the start of each, and let these shared characters be characters 1 and 2.

#### CHARS 1

```
#1. a <1>/ 1. p/ 2. q/
#2. b <1>/ 1. r/ 2. s/
#3. c <1>/ 1. t/ 2. u/
#4. d <1>/ 1. v/ 2. w/
```

#### ITEMS 1

```
#A/ 1<1>,1 2<1>,1 3<1>,1 4<1>,1
#B/ 1<1>,1 2<1>,2 3<1>,1 4<1>,2
```

#### CHARS 2

```
#1. a <2>/ 1. p/ 2. q/
#2. b <2>/ 1. r/ 2. s/
#3. e <2>/ 1. x/ 2. y/
```

#### ITEMS 2

```
#A/ 1<2>,1 2<2>,1 3<2>,1
#B/ 1<2>,1 2<2>,2 3<2>,1
```

#### Step 2.

#### CHARS 2

```
#1. a <2>/ 1. p/ 2. q/
#2. b <2>/ 1. r/ 2. s/
#3. e <2>/ 1. x/ 2. y/
#4. / 1. / 2. / #5. / 1. / 2. /
```

#### Step 3.

#### CHARS 2

```
#1. a <2>/ 1. p/ 2. q/
#2. b <2>/ 1. r/ 2. s/
#3. / 1. / 2. /
#4. / 1. / 2. /
#5. e <2>/ 1. x/ 2. y/
```

#### ITEMS 2

```
#A/ 1<2>,1 2<2>,1 5<2>,1
#B/ 1<2>,1 2<2>,2 5<2>,1
```

#### Step 4.

#### CHARS combined

```
#1. a <1>/ 1. p/ 2. q/
#2. b <1>/ 1. r/ 2. s/
```

```
#3. c <1>/ 1. t/ 2. u/
#4. d <1>/ 1. v/ 2. w/
#5. e <2>/ 1. x/ 2. y/
```

#### Step 5.

#### ITEMS combined

```
#A/ 1<1>,1 2<1>,1 3<1>,1 4<1>,1 1<2>,1
2<2>,1 5<2>,1
#B/ 1<1>,1 2<1>,2 3<1>,1 4<1>,2 1<2>,1
2<2>,2 5<2>,1
```

#### Step 6.

#### ITEMS combined

```
#A/ 1<2>,1 2<2>,1 3<1>,1 4<1>,1 5<2>,1
#B/ 1<2>,1 2<2>,2 3<1>,1 4<1>,2 5<2>,1
```

### Subject: Combining DELTA items

From: Gerhard Rambold

(rambold@botanik.biologie.uni-muenchen.de)

Is there something like an "SUMMARY ITEMS" command? This could probably be useful for persons who intend to use DELTA for the storage of their analysis data, e.g. when they are doing monographical work on species groups in which they have to examine single specimens. Storing such analysis data in DELTA format would allow people to find their taxa e.g. by using the DIST matrix option. After they have found that e.g. ITEM (resp. specimen) no. 4, 7, and 9 belong to one taxon (taxon T), they easily could create a taxon description just by summarizing the ITEM data set.

#### Example:

```
# Specimen 4/ 1,4 2,2 3,6 4,17-19
# Specimen 7/ 1,5 2,2 3,6 4,16-17
# Specimen 9/ 1,5 2,2 3,4 4,16-18
```

```
*SUMMARY ITEMS 4 7 9 ——>
# Taxon T/ 1,4/5 2,2 3,4/6 4,16-19
```

With this method it would also be possible to create ITEMS data sets of genera out of ITEMS data sets of species.

From: Mike Dallwitz, 4 Feb 1995

You can do this with the OUTPUT SUMMARY command in INTKEY. For example,

```
FILE OUTPUT sumitems
```

```
OUTPUT OFF
OUTPUT COMMENT #Taxon T/
OUTPUT SUMMARY (4 7 9) all
```

For combining specimen descriptions into species descriptions, you can use the command SPSUMM (entered at the DOS prompt)<sup>3</sup>. This carries out the above process automatically, but the specimen descriptions to be combined must be contiguous in the ITEMS file, and must all have the same taxon name (apart from comments).

For combining taxon descriptions into descriptions of higher taxa, it is convenient to have a 'character' whose states are the higher taxa. This allows the members of these taxa to be picked out easily. Here is an example from the sample data. We make use of the character:

```
#78. <subfamily>/
1. Pooideae/
2. Bambusoideae/
3. Arundinoideae/
4. Chloridoideae/
5. Panicoideae/
```

In INTKEY, enter the following commands:

```
FILE OUTPUT sfitems
OUTPUT OFF
OUTPUT COMMENT *SHOW Subfamily items
OUTPUT COMMENT *ITEM DESCRIPTIONS
78,1
OUTPUT COMMENT #Pooideae/
OUTPUT SUMMARY remaining all
78,2
OUTPUT COMMENT #Bambusoideae/
OUTPUT SUMMARY remaining all
78,3
OUTPUT COMMENT #Arundinoideae/
OUTPUT SUMMARY remaining all
78,4
OUTPUT COMMENT #Chloridoideae/
OUTPUT SUMMARY remaining all
78,5
OUTPUT COMMENT #Panicoideae/
OUTPUT SUMMARY remaining all
```

These commands (preceded by \*) can be stored in a file and invoked again (by the FILE INPUT command) whenever the data are changed. If you enter the commands interactively, they can be saved via the FILE JOURNAL command for later reuse.

Alternatively, the file can be made up from the character description using a macro in a text editor (I used TED, which is supplied with the DELTA programs, to generate the above example).

The new CONFOR will have much better facilities for this kind of operation. (Had you already guessed that?)

**Subject: INTKEY packages for lichens. Using the same characters at different taxonomic levels**

From: Gerhard Rambold  
(rambold@botanik.biologie.uni-muenchen.de).

To get some idea about how a small interactive species key works under LIAS/INTKEY I have sent you a LIASREV program file for *Rimularia*. This is only a 3-pages revision-module, whereas LIAS includes already 200 pages with key to genera of lichens, various species keys, pictures, data listings etc.

From: Mike Dallwitz, 17 Dec 1994

It is mandatory to include both DELTA.USE and DELTA.REG in your distributions, and to refer your users to DELTA.USE. Among other things, this file also tells how to get updates from the Internet. It would be best to give your users DELTA1@.EXE, which contains all the material relevant to INTKEY, including versions for different languages.

I noticed that the species character list is a greatly reduced version of the genera character list. There would be many advantages in using the same character list for everything [i.e. for the genera and species of ascomycetes]. For example: maintenance of the data would be simplified; you could use the same ICHARS file for all the keys; you could eventually have a single key direct to species (not via genera); you could include genera and species in the same INTKEY data set. This will cause no problems in the species keys, provided that you put \*DISPLAY UNKNOWN OFF in INTKEY.INI.

I suppose that you are aware that many of the species cannot be separated. Don't forget that DIAGNOSE picks this up easily. Perhaps you should mention this to your colleagues who are revising the species keys.

Your final character, which is a text character, has the number of states set to 1 in SPECS. In the next version of CONFOR, we will diagnose this as an error.

From: Gerhard Rambold  
(rambold@botanik.biologie.uni-muenchen.de).

I have thought about your earlier comments about using a single character list, and I have discussed this matter with a number of colleagues who are involved in the LIAS initiative and just preparing character matrices.

Personally I completely agree with you that a uniform character list would be highly desirable to simplify data maintenance and handling. All colleagues who will do the species keys however insisted that they want to be confronted only with their own diagnostic characters which are relevant for their genera or family. I understand well that they do not want to fill out matrices (ITEM files or forms) on the basis of a character list with a huge number of characters only to specify species which are characterizable within their genera or families by a set of ca. 20 characters.

Looking on the ascomycetes there is really such a high diversification of characters in the various groups that all my colleagues agree that one character matrix going down to species level probably would become a colossus. For instance only the character "secondary compounds" would have to include several hundred character states (the various lichen substances) or if I handle it in the way that I treat lichen substances as characters (with two states: "1. absent / 2. present") then there would 600 characters referring only to the chemistry appear on the screen.

If you have no better idea, my intention is now to try to convince my colleagues that they should work out character matrices for the genera which should be uniform at least on family level. Unfortunately the family concept is not yet settled well in most orders of ascomycetes and therefore this procedure will not be realizable up to hundred per cent.

From: Mike Dallwitz, 27 Jan 1995

The new CONFOR will be able to handle very large character sets, restrict the user's view of the characters to a subset, and merge data sets. The technical difficulties that you mentioned will then disappear.

The limit of the current CONFOR running under MS-DOS is around 1000 characters. Even if a combined character list could be kept smaller than this, it would be rather awkward to work with. (Note, however, that

it would NOT be necessary for each contributor to fill in every character for every species.)

The best approach at present would be for each contributor to work with a separate, relatively small character list, but for those lists to include common or 'core' characters at various taxonomic levels. One subset of characters should be included in ALL lists, and other subsets within the lists used by contributors for the various families (or other informal groupings of taxa and/or contributors). That is, each character list would comprise a set of characters for all ascomycetes (agreed among all contributors), a set for a particular family (agreed among the contributors to that family), and a set specific to a particular genus. Such character lists would provide the basis for properly comparative descriptions across the whole group. The data sets could currently be merged (all together or within families) by the method described in an earlier posting, and later by the easier method to be provided in the new CONFOR.

The sample character lists you sent me have some apparently trivial differences which could easily be removed (at any time, but preferably before the data are gathered). For example, here are three characters that apparently record the same information.

Genera:

- #16. lower surface specials (primary thallus)/  
1. absent/  
2. cyphellae/  
3. hypo-, prothallus/  
4. cavernulae/  
5. pseudocyphellae/  
6. rhizines - rhizomorphs/  
7. thallospores/  
8. tomentum - rhizoids - spongiostratum/

Tephromela:

- #24. lower surface specials: <hypo-, prothallus development>/  
1. absent/  
2. hypo-, prothallus present/

Rimularia:

- #10. hypo-, prothallus/  
1. absent/  
2. present/

The first form could be used in all the lists. There are several similar cases.

**Subject: Recording host ranges, etc., in DELTA**

From: Gerhard Rambold  
(rambold@botanik.biologie.uni-muenchen.de).

One problem concerning the lichens is the high number of lichen compounds (more than 600), or, in the nonlichenized parasitic groups, the high number of hosts (hundreds of host genera, thousands of host species).

Even if we concentrate on lichenized Ascomycetes, the matrix should also be suitable for nonlichenized groups, because not the lichens, but the Ascomycetes as a whole are a taxonomical entity.

From: Mike Dallwitz, 2 Feb 1995

There are many difficulties in using these kinds of data in taxonomic databases. There are three basic ways of storing the data in DELTA.

1. As a text character. Advantages: Text characters can accommodate new descriptors (hosts, chemicals, etc.) without any change in the character list. Natural-language descriptions can be made to read exactly as you want. Disadvantages: There is potential for variations in spelling. The data are not very useful for identification (although INTKEY is able to search text characters). They cannot be used for classification.

2. As a single character with each state being a descriptor (e.g. a host species). Advantages: Natural-language descriptions can read fairly well. Use for identification and classification is possible. Disadvantages: The character states must be reordered as new descriptors are added, to maintain a logical or convenient order. (The new CONFOR will have mechanisms for doing this more easily.) Current identification programs (KEY and INTKEY) cannot distinguish between taxa that have different combinations of descriptors/states (e.g. a particular combination of chemicals). Also, if a state is not recorded, there is no way of distinguishing between the case where the taxon definitely does not have the property, and the case where it has not been tested for the property.

3. As a separate character for each descriptor, with states 'present' and 'absent'. Advantages: Identification programs can distinguish between taxa that have different combinations of descriptors (e.g. a particular set of chemicals 'present'). It is possible to

distinguish between 'absent' and 'unknown' (in principle, but see below). Disadvantages: Natural-language descriptions are clumsy. (This might be able to be improved in the new CONFOR). The character list can become very long, and maintenance is clumsy with the current CONFOR.

Whichever recording method is adopted, these characters tend to be unreliable for identification because of the difficulty of getting comprehensive data. Has every lichen species been tested for the presence of each of the 600 compounds? Has every virus been tested for its ability to infect every species of plant or animal? If not, the fact that your lichen specimen has a particular compound, or that your virus infects a particular host, will not eliminate many possibilities even if the data have been recorded by method 3 above. If you want to gamble, you could SET MATCH 0 before using such characters, so as to eliminate the unknown ones.

A useful supplement or alternative to recording hosts at species level is to record them at a higher taxonomic level, say family. One can then be more confident about negative records, and the data are also more useful for classification purposes. This could also be done with the lichen compounds if they have a suitable classification. The new CONFOR will enable convenient and automatic the linkage between characters whose states are different levels of a hierarchy.

**Subject: Regional subsets of taxa in INTKEY**

From: Gerhard Rambold  
(rambold@botanik.biologie.uni-muenchen.de).

Could you please tell me how I should handle the following problem. I would like to include the \*INCLUDE TAXA option, e.g. to use the key only for taxa of certain regions.

However, I do not want to place new items at the end of the ITEMS file but want to place them at the appropriate place in the correct (taxonomic) order. Unfortunately there seems to be no \*NEW ITEM ORDER function in the REORDER option which would transfer the new item at the correct place and would also renumber the \*INCLUDE TAXA numbers.

From: Mike Dallwitz, 19 Jan 1995

There are two ways of doing this. The best is to add a

character or characters with which you code the occurrence of the taxa in the regions. Here is an example from Leslie Watson's Angiosperm-families package.

\*CHARACTER 472

472. <geographical distribution - data complete, intended for keys only>

1. Western and Northern Eurasia, U.S.S.R.
2. Middle East <Turkey to North Africa, and eastwards to Iran>
3. Mediterranean
4. Eastern and Southern Asia <including Japan, China, India>
5. North Africa
6. Central, East tropical and West tropical Africa
7. South Africa
8. Madagascar
9. Malay Archipelago <including New Guinea>
10. Australasia <Australia, New Zealand>
11. Pacific Islands
12. North America
13. Central America
14. West Indies
15. South America
16. Arctic
17. Antarctic

Comprehensively encoded, and intended only for use in identification and for generating useful geographic subsets of the data. Assignments to these pragmatically defined world regions are intended to reflect likelihood of the family being encountered 'in the field', regardless of floristic status. Their reliability has yet to be assessed.

The identification of a specimen from the Antarctic could be started as follows:

\*USE 472,17

472. <geographical distribution - data complete, intended for keys only>

17. Antarctic 4 taxa remain.  
 156. Cyperaceae  
 179. Donatiaceae  
 232. Gramineae 246.  
 Hectorellaceae

The advantage of using this method is that INTKEY's error-tolerance mechanism will apply to this information, so that a specimen from a taxon not previously recorded from the region can still be

correctly identified.

If you want to identify several specimens from the same region, you can 'fix' this character, so that it is not cleared at the start of each identification:

\*SET FIX ON

Characters fixed.

\*RESTART

New identification started (with fixed characters).

\*TAXA REMAINING

156. Cyperaceae  
 179. Donatiaceae  
 232. Gramineae  
 246. Hectorellaceae

If you want to do something other than identification with the regional subset (for example, generating descriptions that are diagnostic within the region) it may be necessary or convenient to exclude all the other taxa:

\*INCLUDE TAXA REMAINING

New identification started.

4 taxa included.

All subsequent operations use only the taxa from the region. The simplest example:

\*TAXA ALL

156. Cyperaceae  
 179. Donatiaceae  
 232. Gramineae  
 246. Hectorellaceae

This method can be applied to any character. For example, with the Angiosperm-families package, you could restrict consideration to monocots, trees, parasitic plants, mangroves, etc.

The second way of restricting operations (including identification) to subsets of taxa is to use the INCLUDE TAXA command with an arbitrary set of taxon numbers or names. For example, restriction to the Antarctic families could be achieved by means of the command

INCLUDE TAXA 156 179 232 246

A better way of specifying a subset is to define a keyword for it, preferably using DEFINE NAMES rather than DEFINE TAXA, because the names are (usually!) more stable than the numbers:



DEFINE NAMES Antarctica  
Cyperaceae  
Donatiaceae  
Gramineae  
Hectorellaceae

The restriction to the Antarctic families could then be done by:

#### INCLUDE TAXA ANTARCTICA

The DEFINE NAMES commands would usually be placed in separate files and invoked by the FILE INPUT command if required (see the Angiosperm-families package for examples). You could place them in INTKEY.INI, so that the keywords would always be available. But if you seldom use the keywords, and their definitions are lengthy, the delay each time the program is started might be irritating.

When used in connection with identification, this method of obtaining regional keys has the disadvantage that the error-tolerance mechanism does not apply - taxa that have been excluded are completely out of contention as possible answers. The advantage is that it is practical to specify regions of narrow interest (such as the campus of a university), which would be impractical to incorporate in the data as a character. Also, any user can define these sets, whereas characters can only be added by the compiler of the data.

The current DELTA programs do not have a NEW ITEM ORDER directive, but this will be provided in the new system.

#### **Subject: INTKEY: direct access to characters which are illustrated**

The following suggestion refers to the Windows version of INTKEY (which will soon be publicly released), but could apply equally well to the MS-DOS version. Mike Dallwitz

From: Joseph H. Kirkbride, Jr.  
(jkirkbri@asrr.arsusda.gov).

It might be good to think about a mechanism which would lead you to only those characters which have images. Being able to see just the images for the characters or taxa without hunting for them could be very useful. When using a new key, it would be much easier to do what I always do, "thumb" through first.

Sometimes the visual gives an immediate answer. For the taxa this is no problem because you can simply select Browsing/Illustrate/Taxa, and you can "thumb" through all the illustrations. If you do not find your organism, then you will look for a character that is easy, one with an illustration. You can similarly "thumb" through the character illustrations by selecting Browsing/Illustrate/Characters, but you cannot use one immediately. You have to close the window, and use one of the options in the Identification menu to select a character. It would be much simpler if you could use a character while browsing through the character illustrations. It seems logical to me because I am always looking for the quickest, easiest way.

From: Mike Dallwitz, 30 Jan 1995

The options in the Browsing menu are for viewing only, and I think it would be confusing to include other functions here. In any case, there are other ways to achieve what you want.

Your suggestion is based on the premise that the illustrated characters will be the easiest or most reliable to use, but this will not necessarily be so. It could be quite the opposite - they might be the ones that are so difficult that an illustration is thought necessary to clarify them. If these characters ARE the easiest, then they should be given high reliabilities, and they will appear near the top of the BEST menus.

As you move down a character-selection list box in BEST or USE, the Images button will be grey if there is no image. (In the DOS version, the function line will indicate the presence of an image.)

If you want a more direct way to use the illustrated characters, you can define a suitable keyword. For example, with the new sample data:

```
*DEFINE CHARACTERS "illustrated" 6 9 10 11 13
```

You can then enter (or select) USE ILLUSTRATED. If DISPLAY IMAGES AUTO is in force, you will then be presented with the character images one by one. You can then select a character state, or reject the character by pressing Esc or Cancel.

Incidentally, identifying a specimen by browsing through the taxon illustrations becomes much more practical with large data sets if most of the taxa are eliminated first by entering some of the attributes of the specimen.

**Subject: Hiding 'inapplicable' character states**

From: Gerhard Rambold  
(rambold@botanik.biologie.uni-muenchen.de)

You mentioned that the new CONFOR will be able to restrict the user's view of the characters to a subset, and merge data sets. Will it also be possible to restrict the user's view of the character states to a subset? E.g.

All genera:

- #16. lower surface specials (primary thallus/)
1. absent/
  2. cyphellae/
  3. hypo-, prothallus/
  4. cavernulae/
  5. pseudocyphellae/
  6. rhizines - rhizomorphs/
  7. thallospores/
  8. tomentum - rhizoids - spongiostratum/

Tephromela:

- #16. lower surface specials (primary thallus/)
1. absent/
  3. hypo-, prothallus/

Miriquidica:

- #16. lower surface specials (primary thallus/)
1. absent/
  3. hypo-, prothallus/
  6. rhizines - rhizomorphs/
  7. thallospores/

Otherwise people would be confronted with character states that do not exist in the genus or even in the order (of the class Ascomycetes) under determination. The consequence would probably be that people then start thinking and trying to find character states in the material that do not exist at all (e.g. there are no cyphellae in the family Lecanoraceae).

From: Mike Dallwitz, 2 Feb 1995

This could be done fairly easily, but we have no plans to do so. A closely related question arises during interactive identification: if some character states do not occur in the taxa remaining at some stage of the identification, should these states be marked (or hidden or made unavailable)? Some programs, e.g. Pankhurst's ONLINE, do this. I am against it because it biases the user's response towards agreement with previously entered information. If some of the previous

information was wrong, the new information is more likely to be wrong too, which makes it more difficult to recover, whether by manually correcting the errors or by INTKEY's error-tolerance mechanism.

Because it is generally more efficient to identify directly to species (rather than first to family, then to genus, then to species), you may eventually want to combine all your data into a single dataset. In that case, you will be in exactly the situation discussed in the previous paragraph. Even if you keep the datasets separate, the same argument applies, because the wrong dataset may be being used because of an earlier mistake in identification to (say) genus. Your concern that people may be confused by the 'extra' states is probably mainly for people who are unfamiliar with the group. But these are actually the ones who would benefit most by the inclusion of those states!

**Subject: Using DELTA to produce non English descriptions and keys**

From: Joseph H. Kirkbride, Jr.  
<jkirkbri@asrr.arsusda.gov> 15 Dec 1994

In October and November of this year, I taught a course on the DELTA software system in the Departamento de Botânica, Universidade de Brasília, Brasília, Brazil, and gave various demonstrations of the DELTA software. The course and demonstrations were presented in Portuguese, and the attendees were interested in producing products, descriptions, keys, etc., in Portuguese.

Consequently I placed the command \*VOCABULARY in the SPECS file followed by the vocabulary supplied by Dallwitz in the file VOCABP. CONFOR worked fine, and I had no trouble producing descriptions, keys, or files for INTKEY. When I tried to demonstrate DELFOR using TIDY and REORDER, I could not get it to work because of approximately 20 error messages associated with the command \*VOCABULARY and the Portuguese vocabulary. When I removed the command \*VOCABULARY and the Portuguese vocabulary from SPECS, DELFOR functioned to tidy or to reorder the data files.

I contacted Dallwitz in Australia by e-mail and described what had happened. He informed me that to use a non English vocabulary, i.e., Portuguese, to put into TONAT &TOKEY the following command line:

```
*INPUT DELTA FILE VOCABP
```

With that command line in TONAT and TOKEY, I am able to use CONFOR to produce Portuguese descriptions and keys and to call DELFOR to tidy and reorder the data.

From: Mike Dallwitz, 16 Dec 1994

The problem with the VOCAB directive, mentioned by Joe Kirkbride in an earlier posting, is a DELFOR bug. It will be fixed in the next release. In the meantime, please use the workaround given in Joe's posting.

**Subject: Using INTKEY over the Internet**

From: Anthony Brach <brach@huh.harvard.edu>

Is it possible (or in the future) for users of WWW to utilize DELTA INTKEY or a modification thereof at a particular site of the WEB ? Since users can presently access checklists, descriptions, non-interactive keys, and descriptions, thinking about the potential for interactive keys of floras online.

From: Mike Dallwitz, 16 Dec 1994

Version 2 of INTKEY is available for Unix (SunOS), and can be used across the Internet. For an example, see

Christiansen, K. A., Bellinger, P. F., and da Gama, M. M. (1990). Computer assisted identification of specimens of *Pseudosinella* (Collembola Entomobryidae). *Rev. Ecol. Biol. Sol* 27, 231-246.

To try this system:

telnet ac.grin.edu  
Username: BUGS  
Password: DELTA

Version 2 of INTKEY has only a command-line interface (though lists of commands can be displayed either manually or automatically). We hope eventually to port the Windows version of INTKEY (version 4) to other GUI's, including Mac and Unix. I don't know whether either version could be accessed via WWW.

**Subject: DELTA and Pentium**

From: Anthony Brach <brach@huh.harvard.edu>

As you've probably heard, Intel Pentium's are miscalculating when applications utilize the processor for floating point divide calculations. Can you please let us know if DELTA uses floating point divide calculations or if there is a program that compensates for the error?

Reply from Mike Dallwitz. 15 December 1994

Many programs use the floating-point processor if available, and otherwise use a subroutine library which emulates that processor. The DELTA programs use a different kind of floating-point emulation, called ALTMATH, which works faster on computers that don't have a floating-point processor; the floating-point processor is never used. This decision was taken when floating-point processors were rare.

We probably should consider making our programs use the floating-point processor, but this is obviously not an opportune time to do so. The only programs which would benefit substantially are DIST and TYPSET.

**Subject: Screen capture**

From: Marie-Claude Lariviere  
<LariviereM@landcare.cri.nz>

I would like to know if you can recommend a program to capture what's on my screen when I use INTKEY with my own data so as to send the image to a file and retrieve it later for use as a Figure in my paper.

From: Mike Dallwitz, 19 Jan 1995

I use GRABBER<sup>4</sup> to capture DOS screens, and SuperClip for Windows. They are on the DELTA servers as graphics/grab394.exe and graphics/sclip15@.exe. GRABBER superimposes its logo on the captured screens unless you register it. There are many other screen-capture programs around on the Internet.

From: Eric J. Gouda <E.J.Gouda@CC.RUU.NL>, 19 Jan 1995

If you run DOS programs in a Windows box you can easily use the standard edit function (Mark and Paste) of Windows and insert them in any editor also running.

From: Mike Dallwitz, 20 Jan 1995

This method captures the window as text. If this text is pasted into a Windows text editor, the box-drawing characters (used in INTKEY menus and messages) are converted to +, |, and -. If it is pasted into the DOS EDIT program running in a window, the box-drawing characters cause the editor to behave erratically.

To capture a Windows screen or the active window (including a DOS application running in a Window),

you can press PrintScreen or Alt+PrintScreen to copy the screen or window to the clipboard, and then paste it into most paint programs (but it won't work correctly with the standard Windows Paintbrush accessory). SuperClip uses this method, but is more convenient than a normal paint program for capturing long sequences of screens, because it pops up automatically whenever the clipboard changes. It also has a convenient cropping mechanism (you will probably want to crop DOS screens captured in this way, to remove the Windows header and border).

GRABBER captures DOS text or graphics screens. Text screens can be converted to image files (with a choice of 4 fonts) or text files (with or without ANSI escape sequences).

From: Marie-Claude Lariviere, 20 Jan 1995

Dear all, Following our discussion on screen capture .... Thanks Mike, Robin and Eric for your inputs. This morning I didn't know how to capture INTKEY screens. Since then, I have got myself a copy of GRABBER from the DELTA server in Australia.

GRABBER is great ... it's easy to use and real quick. Screens are saved to an .exe file during a INTKEY

session simply by pushing 2 keys. The advantage of .exe files is that one can look at the saved screens simply by typing the name of the .exe file at the DOS prompt. In addition, a simple DOS command is required (provided with GRABBER) to convert the .exe file to a bitmap format file, for example .pcx, which can then be picked up by most 'paint' programs.

I recommend the use of GRABBER. If like me you're not an expert in computer imaging and so forth, this is a quick and easy program to use.

Marie-Claude Lariviere New Zealand Arthropod Collection, Landcare Research, Private Bag 92170, Auckland, New Zealand Tel 64 9 849 3660, Fax 64 9 849 7093, Email LariviereM@landcare.cri.nz

#### Notes

1. For further details see R. Cowan, *DELTA Newsletter* 10: 11 (1994).
2. The DELTA Menu System for Windows (DMSWin) of N.S. Lander (*DELTA Newsletter* 8: 10-11 (1993)) provides the same functionality.
3. SPSUMM is described by T.D. Macfarlane, *DELTA Newsletter* 8: 12 (1993).
4. PaintshopPro is another widely used shareware program for screen capture.



#### RECENT PUBLICATIONS

ALICE Software News 3 (Autumn 94).

Contents:

Introduction

ALICE System 2 Developments (Latest versions; ACE; AMIE Version 2.3.2; SAM Version 2.2.8; New releases of other programs; Third-party add-ons; Work in progress; ALICE for Windows)

ALICE System 3 (Evolutionary database development; New functionality; The future)

Documentation (Progress and distribution; Translation; The next steps)

Data exchange (Links with specimen based database systems; Importing DELTA format; Exporting to cladistic programs)

New licensing policy (Purchasing rights; Annual user licences; Fees; Rights of licensees)

Contact addresses:

Dr Bob Allkin (Alice Software), c/o Royal Botanic Gardens Kew, Richmond, Surrey TW9 3AB, UK. Email asoft@cix.compulink.co.uk and r.allkin@rbgkew.org.uk.

Peter Winfield, Alice Software, 28 Mardale Crescent, Edinburgh EH10 5AG, Scotland, UK. Email winfield@sasa.gov.uk.

Loizeau, P.-A. (1994). Les Aquifoliaceae peruviennees. (Elements pour une revision des Aquifoliaceae neotropicales). *Boissiera* 48. 306pp.

From the English abstract: The information is stored in electronic form, using data bases for the specimina visa, nomenclature and bibliography, or the DELTA system, a descriptive language for taxonomy which allows to make interactive determinations and to produce descriptions in a "natural language" or, automatically, classical keys of determination.

Ding Hou (1994). Studies in Malesian Caesalpiniaceae (Leguminosae). I. The genera *Acrocarpus*, *Afzelia*, *Copaifera*, and *Intsia*. *Blumea* 38: 313-330.

Kirkbride, J.H., Jr. (1994). Revision of *Cucumella* (Cucurbitaceae, Cucurbitoideae, Melothrieae, Cucumerinae). *Brittonia* 46: 161-186.

Watson, L & Dallwitz, M.J. (1994). *The Grass Genera of the World*. Revised Edition. (CAB International: Wallingford.)

**Release of INTKEY for Windows**

**T.D. Macfarlane**

INTKEY for Windows (INTKEY4) was released this month, following two years of development, including testing and consultation with users. It operates under Microsoft Windows 3.1, supports images and sounds, and employs the same initialization file and CONFOR-produced data files as the DOS version 3. Likewise the command/function set is almost identical.

Toni Paine presented a previous description of the program (DELTA Newsletter 10: 5-6, April 1994). Although substantially similar to the development version described there, some changes have been made, partly as a result of user feedback. At start-up there is a choice of three modes, Simplified, Automatic and Manual, so that users can match their needs and level of familiarity to an appropriate interface. The top level menubar has reworded menu titles. The character window has notes and image buttons which are greyed when notes or images are not available, and a search facility for finding strings in character features. Overlay of text on character images is supported, including active areas which allow direct selection of states in identification mode.

INTKEY for Windows is available by gopher or anonymous ftp from the following Internet hosts.

muse.bio.cornell.edu (directory: /pub/delta)  
spider.ento.csiro.au (directory: /delta).

People who are already registered users of INTKEY will not have to pay an additional registration fee.

**DELTA NEWSLETTER BACK ISSUES**

Limited back issues of *DELTA Newsletter* numbers 1 - 10 are now available. These issues contain much useful information and advice and provide an invaluable reference to any biologist interested in the area of computer-aided taxonomy. They can be ordered individually or as a set. Please indicate which issues you require and how many of each. Each issue is a fixed price of \$5.00 each, which includes postage and handling.

Please note that extra-Australian personal cheques cannot be dealt with; instead make International Money Orders payable to **The DELTA Newsletter**.

Name \_\_\_\_\_

Address \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

Country \_\_\_\_\_

Email address \_\_\_\_\_

Please send me the following DELTA Newsletter back issues. (Tick applicable box for each Newsletter requested and indicate if multiple copies are required.)

1 x\_\_\_\_ 2 x\_\_\_\_ 3 x\_\_\_\_ 4 x\_\_\_\_ 5 x\_\_\_\_.  
6 x\_\_\_\_ 7 x\_\_\_\_ 8 x\_\_\_\_ 9 x\_\_\_\_ 10 x\_\_\_\_.

Total number of back issues: \_\_\_\_ @ \$US5.00

Total payment enclosed: \$US\_\_\_\_\_.

Signed \_\_\_\_\_ Date \_\_\_\_\_

Send to: The Editors  
DELTA Newsletter  
Western Australian Herbarium  
PO Box 104  
Como WA 6152  
AUSTRALIA



## About the DELTA Newsletter

A communications medium for botanical and zoological taxonomists interested in descriptive databases.

| <b>Topics</b>                  | <b>Features</b>                  | <b>Closing dates</b> March & September |
|--------------------------------|----------------------------------|--|
| Computer programs for taxonomy | Articles                         | <b>Submissions welcome.</b>            |
| Data formats                   | Program releases and updates     | Post, fax or email to:                 |
| Data interchange standards     | Spotlight on DELTA features      |  |
| Data capture                   | Technical tips                   | The Editors                            |
| Data analysis                  | Database advertisements          | DELTA Newsletter                       |
| Database design                | Meeting notices and reports      | Western Australian Herbarium           |
| Description printing           | Publication and database reviews | PO Box 104                             |
| Expert systems                 | and notices                      | Como WA 6152                           |
| Information retrieval          | Image issues                     | AUSTRALIA                              |
| Interactive identification     | Letters                          |  |
| Keymaking                      | Personal profiles                | Phone +61 9 334 0500                   |
| Mapping systems                |                                  | Fax +61 9 334 0515                     |
| Taxonomic characters           | <b>Frequency</b>                 | Email (via M. Dallwitz) Internet       |
|                                | Twice yearly, April and October  | md@ento.csiro.au                       |

### SUBSCRIBE TO THE DELTA NEWSLETTER

To subscribe to the DELTA Newsletter, please complete the form below and return to the editors.

Please put me on the mailing list for future issues of the DELTA Newsletter.

Name \_\_\_\_\_

Address \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

Country \_\_\_\_\_

Email address \_\_\_\_\_

Research interest / Taxonomic speciality \_\_\_\_\_

\_\_\_\_\_

Send to: The Editors  
DELTA Newsletter  
Western Australian Herbarium  
PO Box 104  
Como WA 6152  
AUSTRALIA

Signed \_\_\_\_\_ Date \_\_\_\_\_



