PANKEY

Programs for the identification and description of plants of animals

R.J.Pankhurst
Taxonomic Systems
Edinburgh

Exeter Software

47 Route 25A, Suite 2, Setauket, New York 11733-2870

PANKEY is a package of programs for problems of identification or diagnosis. It is usually applied to the naming of plants or animals, but it can also be used for medical diagnosis and for problems outside the life sciences. Program output of diagnostic keys and descriptions of species will typically be used in taxonomic publications such as monographs, Floras and Faunas. Also included is an expert identification program which is used interactively, with images, in order to name unknown specimens. Data for **PANKEY** is stored in the DELTA format, which is an international standard.

The **PANKEY** programs include:

- construction of diagnostic keys, both automatically and interactively
- construction and printing of species descriptions
- expert interactive identification, with colour images
- identification by comparison (matching)
- character analysis
- conversion to other formats, for clustering or cladistics
- DEDIT, a special purpose editor for DELTA data (comes free of charge)

PANKEY as described here is a development package, so that you can set up and modify your own data sets. You are encouraged to share and distribute the data files that are created, free of charge. Special licences are available for the distribution of runtime only **PANKEY** programs. **PANKEY** is particularly suitable for teaching, in conjunction with the textbook "Practical Taxonomic Computing" by R.J.Pankhurst (Cambridge University Press, 1991).

PANKEY comes with a 150 page user manual, and is provided with online help. **PANKEY** system requirements are very modest. This is deliberate, so that it can be used in all parts of the world. The minimum machine specification is:

- 2.4 Mb of hard disk
- 640K RAM (but extended memory is better)
- 286 processor or above
- DOS 5 or above (operating system)
- monochrome VGA graphics (although colour is better)

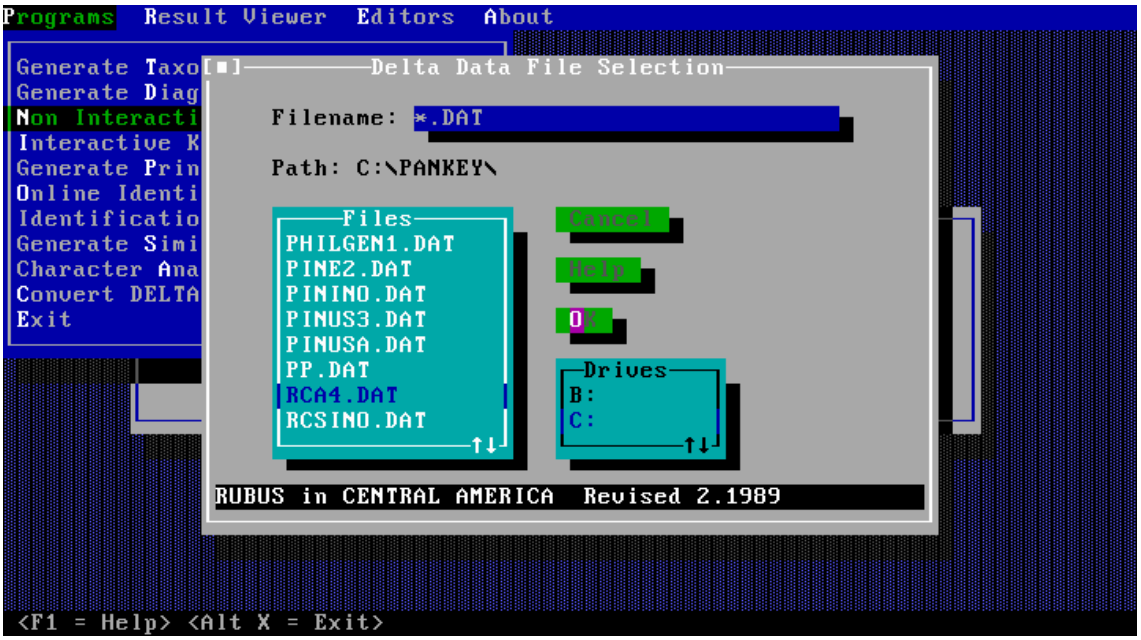
R.J.Pankhurst *Exeter Software*

EMAIL: *Richard@rbge.org.uk* *Sales@ExeterSoftware.com*
INTERNET: *www.ExeterSoftware.com*

Phone: *+44 -131-248-2914* *+1-631-689-7838*
Fax: *+44-131-248-2901* *+1-631-689-0103*

PANKEY startup screen

This shows the selection of one of the 10 program options and the selection of an input DELTA data file.



Key-generation program, automatic

The only interaction needed is to specify the data file**.** The following key was generated from a DELTA data set for *Jurinea* (see Appendix)

```
1    Sterile rosettes present.                                          2
 2    Outer involucral bracts patent, or recurved.                      3
  3    Cauline leaves auriculate, pappus 0.8 times achene.
                                                        11.J.polyclonos
  3    Cauline leaves without auricles, pappus 1.1 to 1.2 times
       achene.                                                          4
   4    Cauline leaves amplexicaul.                     12.J.ledebourii
   4    Cauline leaves not amplexicaul.                                 5
    5    Achene 1.0 to 2.0 mm.                         14.J.glycacantha
    5    Achene 3.0 to 5.0 mm.                              10.J.mollis
 2    Outer involucral bracts erect.                                    6
  6    Upper surface of leaves white, or grey, basal leaves
       arachnoid-tomentose above.                                       7
   7    Basal leaves entire, leaf margins revolute, capitula
        hemispherical, corona of achene inconspicuous, pappus 1.2 to
        1.4 times achene.                               7.J.kirghisorum
   7    Basal leaves pinnatifid, leaf margins plane, capitula
        obconical, corona of achene conspicuous, pappus 3.0 to 4.0
        times achene.                                        4.J.pinnata
  6    Upper surface of leaves green, basal leaves subglabrous above,
       or setose above.                                                 8
   8    Basal leaves setose above, capitula obconical, corona of
        achene conspicuous, pappus 3.0 to 4.0 times achene.
                                                  3.J.tzar-ferdinandii
   8    Basal leaves subglabrous above, capitula subglobose, or
        hemispherical, corona of achene absent, or inconspicuous,
        pappus 1.1 to 2.0 times achene.                                 9
    9    Capitula subglobose, corona of achene absent.  17.J.fontqueri
    9    Capitula hemispherical, corona of achene inconspicuous.
                                                      13.J.consanguinea
1    Sterile rosettes absent.                                          10
10    Leaf margins revolute.                                           11
 11    Basal leaves setose above.                                       12
  12    Pappus 1.5 to 2.0 times achene.             2.J.stoechadifolia
  12    Pappus 3.0 to 4.0 times achene.           3.J.tzar-ferdinandii
 11    Basal leaves arachnoid-tomentose above.                          13
  13    Capitula cylindrical, achene glabrous.        1.J.linearifolia
  13    Capitula obconical, achene hairy.                               14
   14    Outer involucral bracts erect, distal part of bracts purple.
                                                        16.J.taygetea
   14    Outer involucral bracts patent, or recurved, distal part of
         bracts green.                                   15.J.humilis
10    Leaf margins plane.                                              15
 15    Stem woody at base.                                              16
  16    Achene 6.0 to 7.0 mm, corona of achene inconspicuous, pappus
        1.5 times achene.                                6.J.albicaulis
  16    Achene 3.0 to 4.5 mm, corona of achene conspicuous, pappus
        2.5 to 4.0 times achene.                                        17
   17    Basal leaves entire, capitula cylindrical.   1.J.linearifolia
   17    Basal leaves pinnatifid, capitula obconical.      4.J.pinnata
 15    Stem herbaceous.                                                 18
  18    Upper surface of leaves white, or grey, basal leaves
        arachnoid-tomentose above, capitula obconical, outer
```

```
      involucral bracts coriaceous.                    5.J.tanaitica
 18   Upper surface of leaves green, basal leaves subglabrous above,
      capitula subglobose, outer involucral bracts herbaceous.    19
 19    Cauline leaves without auricles, outer involucral bracts
       erect, achenes obpyramidal.                     8.J.cyanoides
 19    Cauline leaves auriculate, outer involucral bracts recurved,
       achenes subcylindrical.                       9.J.ewersmanii
```

This key can be transferred to the next program, interactive key construction, for detailed editing and unlimited improvement, only provided that the DELTA data file is unchanged.

For further reading and explanation of the program, see:

PANKHURST, R.J. (1970) A computer program for the generation of diagnostic keys. Computer Journal, 12:145-151.

PANKHURST, R.J. (1970) Key generation by computer. Nature 227: 1269-1270.

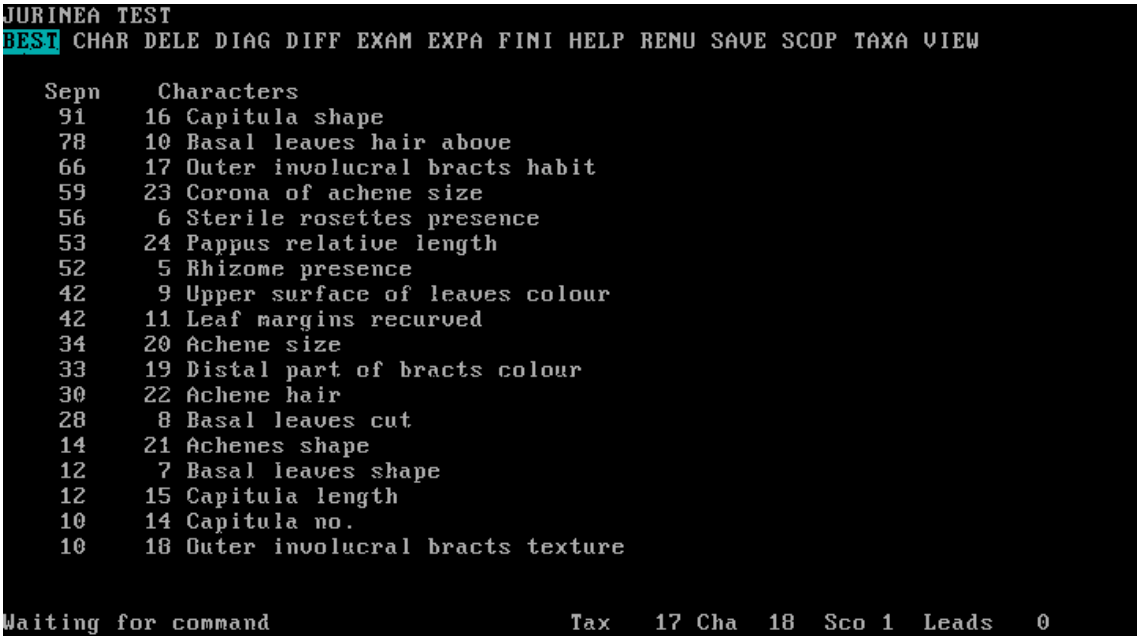
PANKHURST, R.J. & WALTERS, S.M. (1971) Key generation by computer, in Data processing in Biology and Geology, ed. J.L.Cutbill, for Systematics Association, Academic Press: 189- 203.

PANKHURST, R.J. (1971) Botanical keys generated by computer. Watsonia 8:357-368.

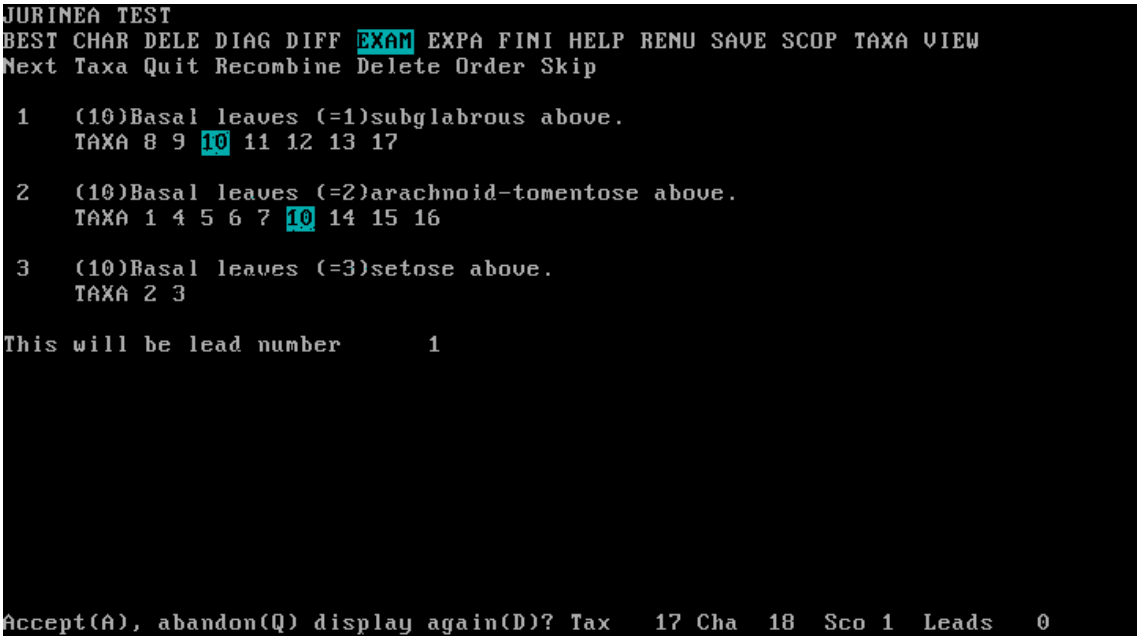
PANKHURST, R.J. (1991) “Practical Taxonomic Computing” Cambridge University Press. A text book of taxonomic computing.

Key-generation program, interactive

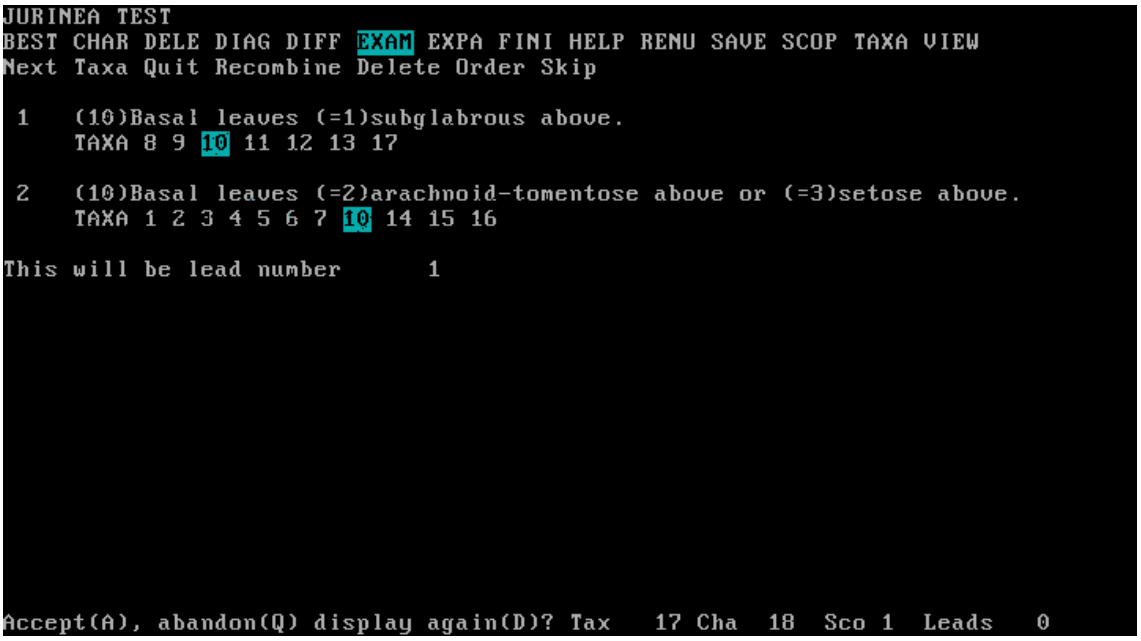
Screen 1: Using the BEST command to find which characters contain the most information for key construction. The same command appears in the ONLIN7 program.



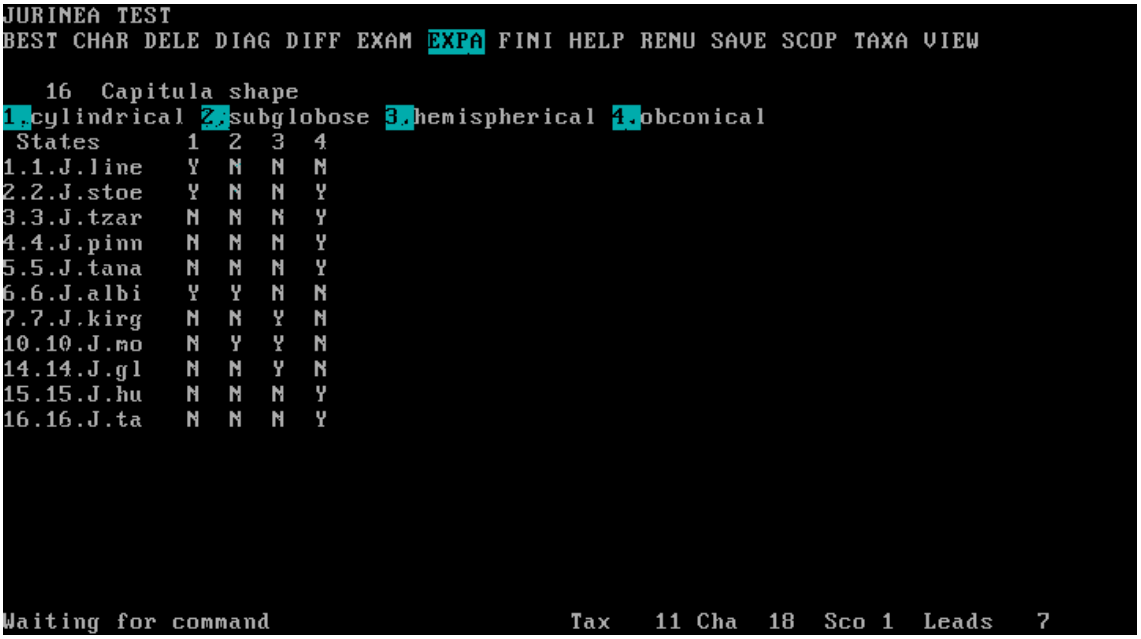
Screen 2: Using the EXAM command to examine the leads of a key which would result from using the character 10 ‘Basal leaves hair above’. Duplicated taxa due to variable characters are highlighted in blue.



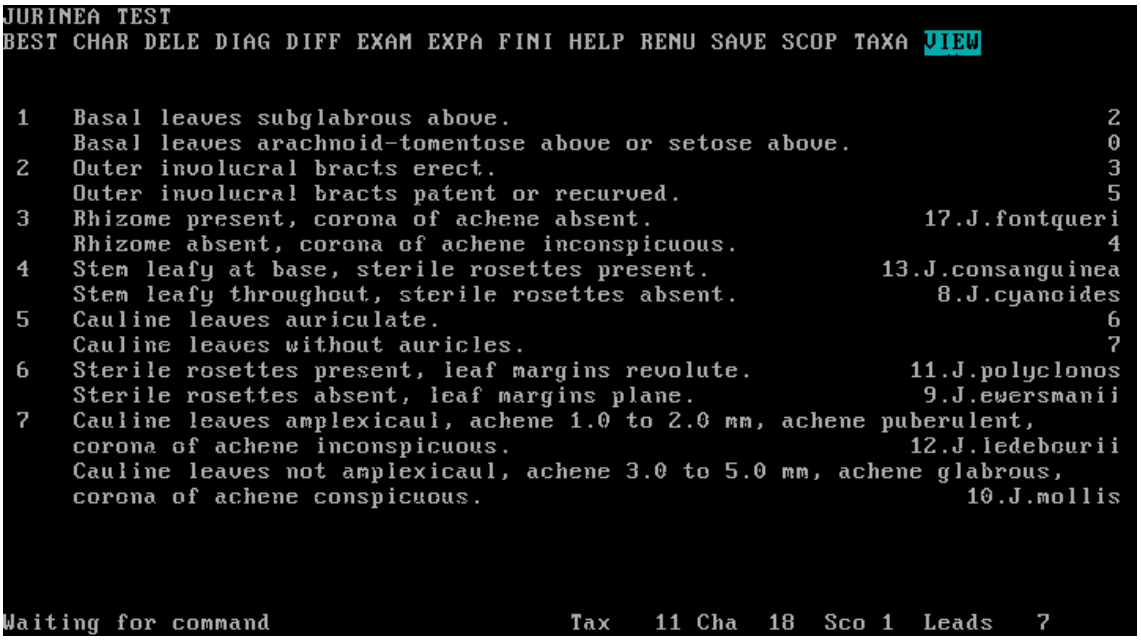
Screen 3: Since this key has three leads, and a dichotomous key would be preferred, the two leads for 'hairiness' are combined.



Screen 4: Expanding the 'Capitulum shape' character in order to see which states are involved and how they are distributed.



Screen 5: Viewing the partially constructed key. Although this is shown in the 'parallel' style without indentation, other key styles are available at output time.

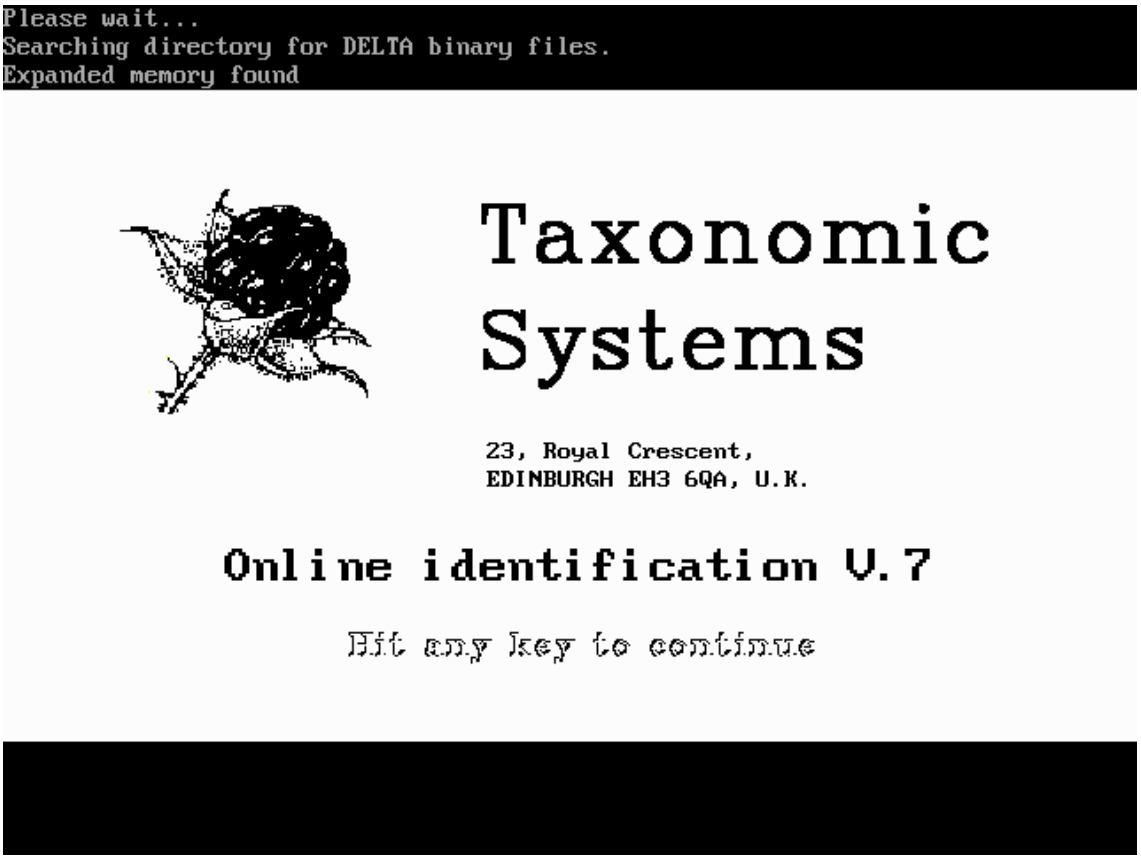


Further reading:

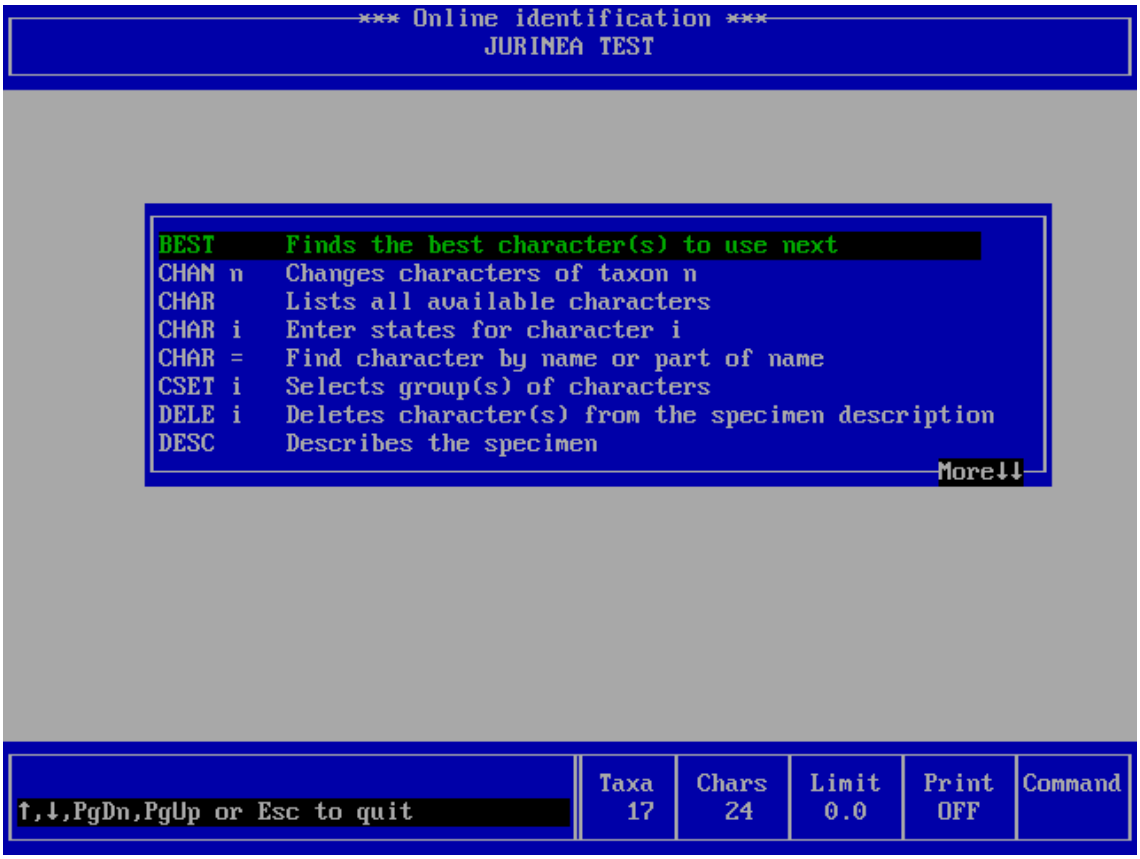
PANKHURST, R.J. (1988) An interactive program for the construction of identification keys. Taxon 37(3): 747-755.

On-line identification

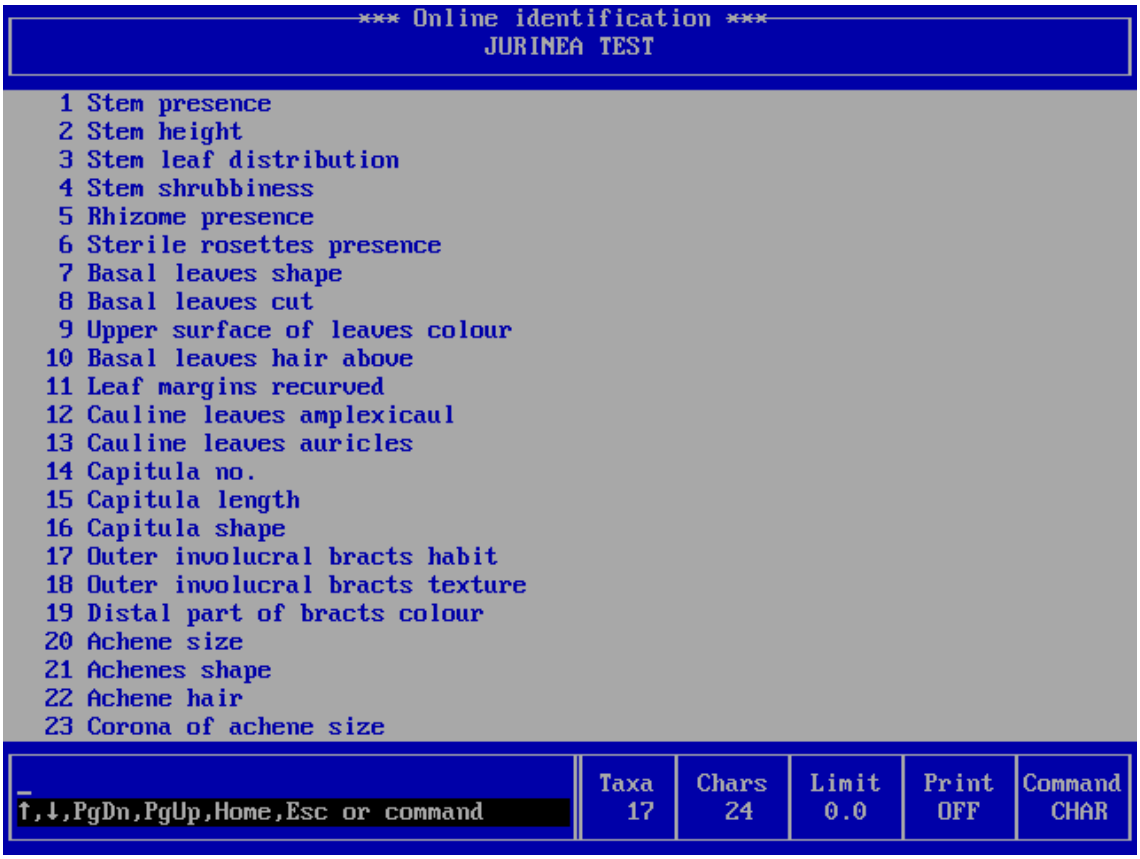
Screen 1: Startup screen



Screen 2: Menu of available commands

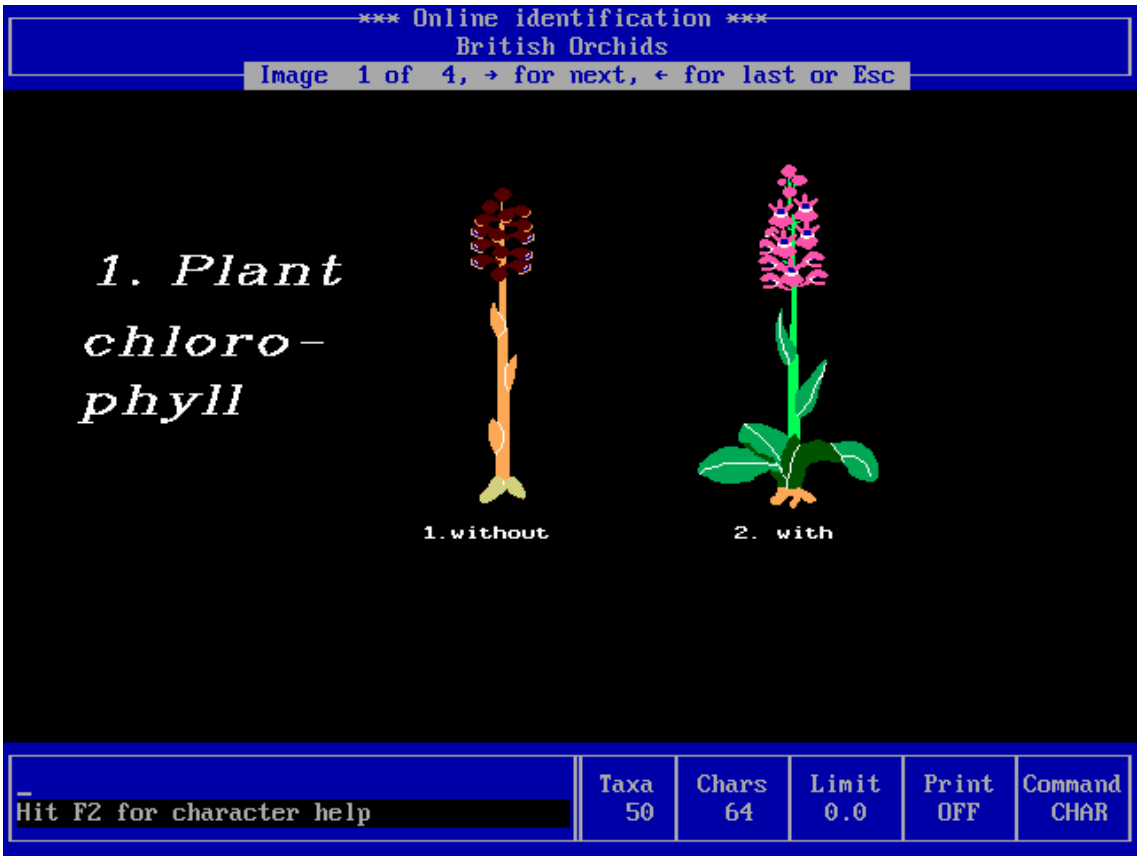


Screen 3: Listing of all available characters. You can just pick the character that you feel like using and go ahead.



Screen 4: BEST command, showing character information values, with weighting.
This is much better than just choosing characters from the list, unless you are an expert in the group, since you will advance much more rapidly towards a solution. You can also use the DIAG command, to get characters for a specific taxon, and the DIFF command, to see differences between two taxa, when you think you might have one of them, but are not sure which. It is also possible to set and to vary, the number of 'wrong' characters that you are willing to allow, with the LIMI command. If the LIMIt is zero, then the specimen has to agree exactly with the target taxon, or if set to 1, then one difference is allowed, and so on.

Screen 5: When choosing a state for a character, illustrations can be provided. This image comes from the classic key for British orchids, showing the difference between a saprophytic and an autotrophic species. Illustrations can also be attached to the TAXA command, for viewing once a tentative identification has been reached.



For further reading:

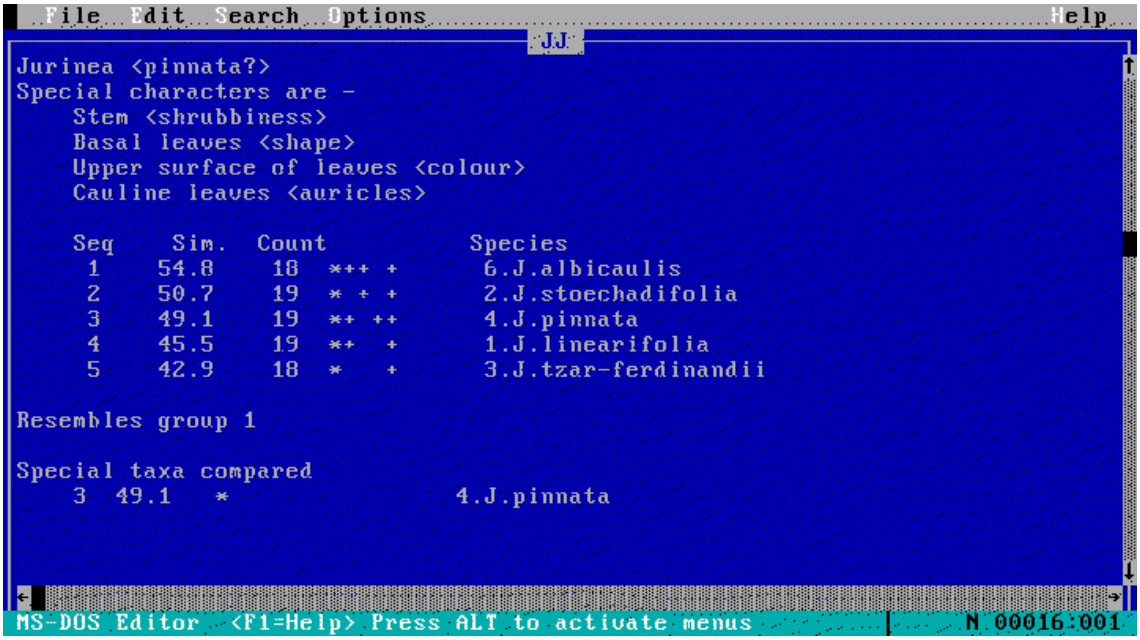
PANKHURST, R.J. & AITCHISON, R.R. (1975) An on-line identification program. In "Biological Identification with Computers", ed. R.Pankhurst for the Systematics Association, Academic Press: 181-194.

PANKHURST, R.J. (1989) A computer program with colour graphics to identify orchids. Orchid Review 97(1144): 53-55,67.

Identification by Matching

This program does not carry out the step by step elimination of the taxa by characters as in a key or an online program but calculates a similarity between a complete description of the specimen and each of the taxa

Screen 1: Shows the highest scoring taxa in the genus *Jurinea*, with plus signs for the 'special' characters that have been pre-selected. The correct answer is not necessarily the taxon with the highest score.



Further reading:

PANKHURST, R.J. (1975) Identification by matching. In "Biological Identification with Computers", ed. R.Pankhurst for the Systematics Association, Academic Press: 79-92.

Description printing

The following text was generated using the standard PANKEY DELTA data example '*Jurinea*' and shows the generalised description using constant or nearly constant characters (qualified by 'usually') in the first section of the genus, and then a description of the first species 'linearifolia' with characters set out in paragraphs with (optionally) their DELTA character numbers left in, as an aid to proofreading the correctness of the DELTA data.

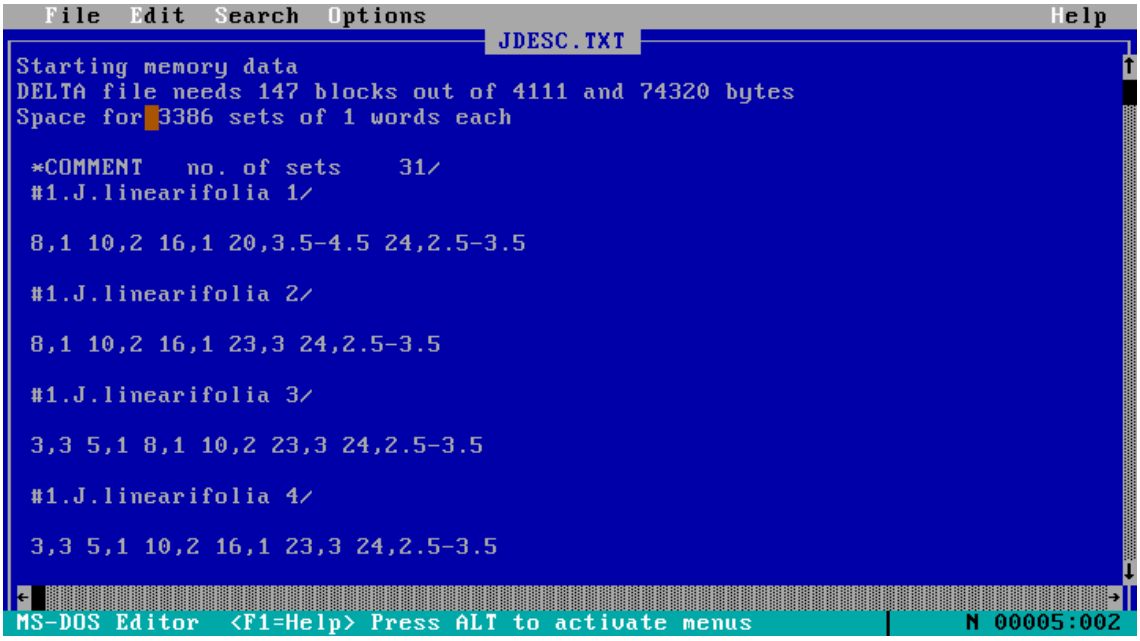
```
JURINEA PART 1
     (5) Rhizome usually absent. (1) Stem present, (2) 4.0 to 80.0 cm, (3)
usually leafy throughout. (12) Cauline leaves usually not amplexicaul,
(13) usually without auricles. (14) Capitula 1 to 20, (15) 0.5 to 7.0 cm.
(17) Outer involucral bracts usually erect. (20) Achene 3.0 to 7.0 mm,
(22) usually glabrous. (24) Pappus 0.8 to 4.0 times achene.


1.J.linearifolia
     (5) Rhizome absent. (6) Sterile rosettes absent. (1) Stem present,
(4) woody at base, (2) 12.0 to 40.0 cm, (3) leafy throughout.
     (7) Basal leaves linear-lanceolate or lanceolate, (8) entire, (10)
arachnoid-tomentose above. (11) Leaf margins plane or revolute. (12)
Cauline leaves not amplexicaul, (13) without auricles.
     (15) Capitula 0.5 to 1.8 cm, (16) cylindrical. (17) Outer involucral
bracts erect, (18) herbaceous. (19) Distal part of bracts pink, reddish,
or purple.
     (20) Achene 3.5 to 4.5 mm, (21) tetragonal, (22) glabrous. (23)
Corona of achene conspicuous. (24) Pappus 2.5 to 3.5 times achene.



2.J.stoechadifolia

......
```

Further reading:

PANKHURST, R.J. (1978) The printing of taxonomic descriptions by computer. Taxon 27:65-68.

Diagnostic descriptions

Screen 1: The program has been set to find all diagnostic character sets of minimum size with the LIMIT set to 1, so that the 31 different character sets will each distinguish *Jurinea linearifolia* from all other taxa by at least 2 characters



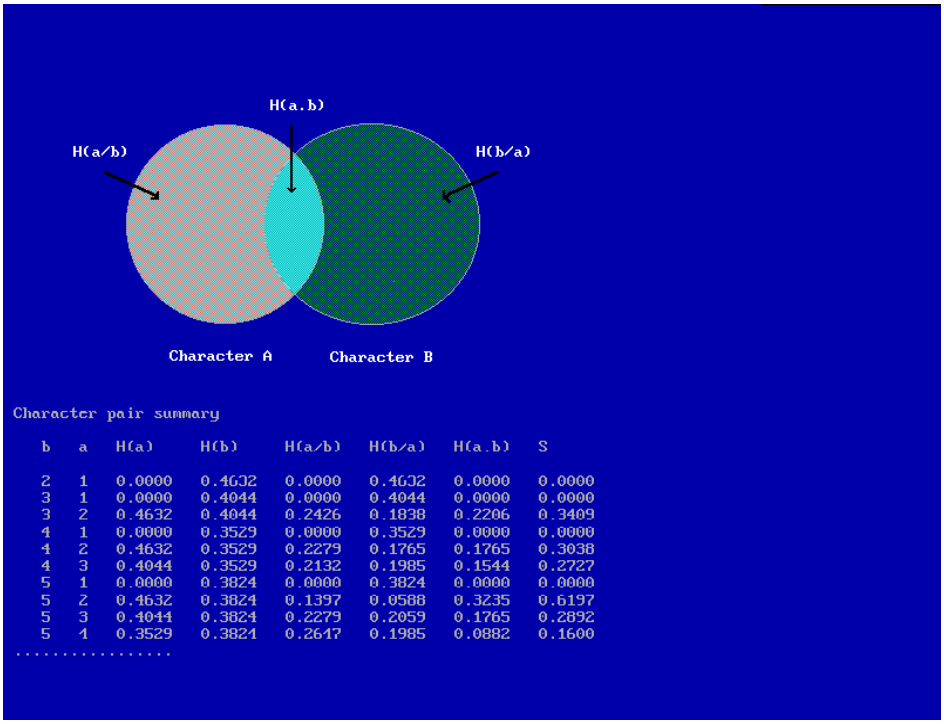
Further reading:

PANKHURST, R.J. (1983) An improved algorithm for finding diagnostic taxonomic descriptions. Mathematical Biosciences 65: 209-218.

Character analysis

This program provides a means to assess the correlations between characters. A check is advisable because a DELTA character set might contain redundancy. This cannot be calculated in the statistical sense because we do not have strictly quantitative variables, but instead a mixture of (mainly) qualitative characters and a few quantitative ones. A formula based on the information statistic is used, so that the value S in the diagram represents the common area of variation between two characters, illustrated as the overlapping area between two characters, whose variation is represented by circles. If S approaches unity, then the two characters might be

1) two ways of saying the same thing i.e one is redundant, or
2) a sign of a meaningful taxonomic grouping, or an indication of synapomorphy or symplesiomorphy



Character pair summary

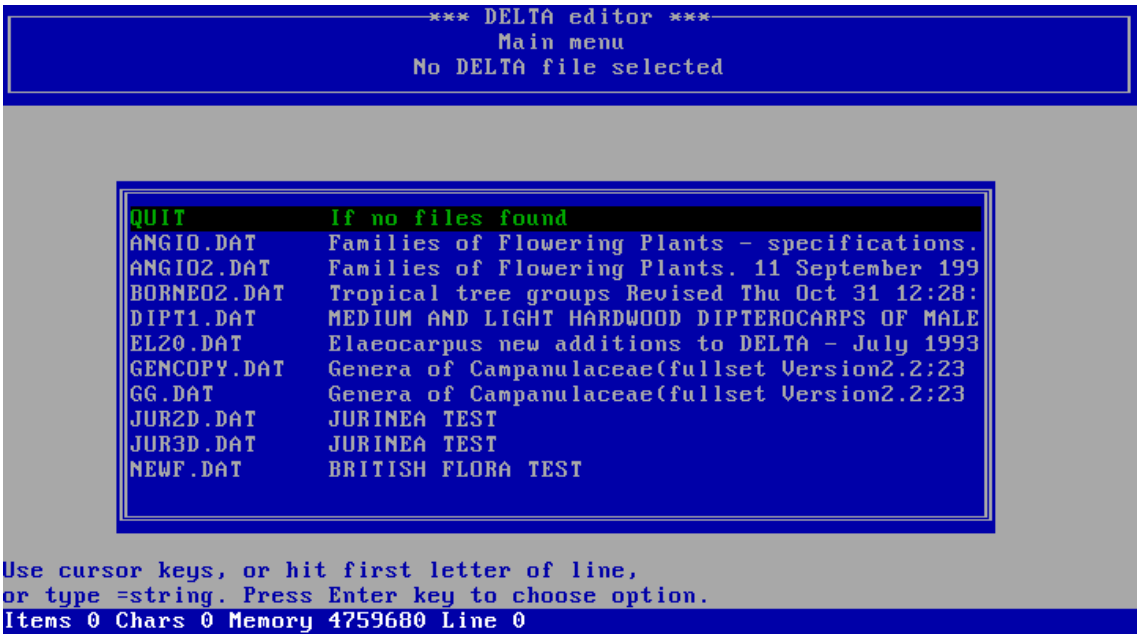
b	a	H(a)	H(b)	H(a/b)	H(b/a)	H(a.b)	S
2	1	0.0000	0.4632	0.0000	0.4632	0.0000	0.0000
3	1	0.0000	0.4044	0.0000	0.4044	0.0000	0.0000
3	2	0.4632	0.4044	0.2426	0.1838	0.2206	0.3409
4	1	0.0000	0.3529	0.0000	0.3529	0.0000	0.0000
4	2	0.4632	0.3529	0.2279	0.1765	0.1765	0.3038
4	3	0.4044	0.3529	0.2132	0.1985	0.1544	0.2727
5	1	0.0000	0.3824	0.0000	0.3824	0.0000	0.0000
5	2	0.4632	0.3824	0.1397	0.0588	0.3235	0.6197
5	3	0.4044	0.3824	0.2279	0.2059	0.1765	0.2892
5	1	0.3529	0.3824	0.2647	0.1985	0.0882	0.1600

.....

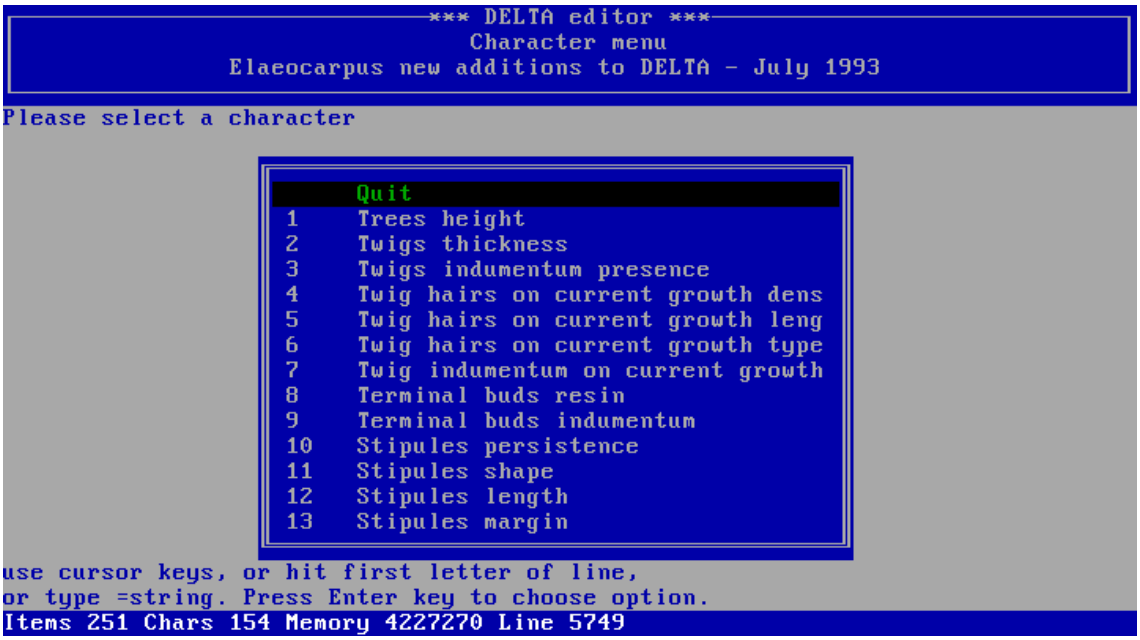
DELTA editor

The PANKEY package provides a special purpose editor for DELTA data, called DEDIT for short. Think of it as a text editor for DELTA, in the same way as a word processor is a system for writing letters.

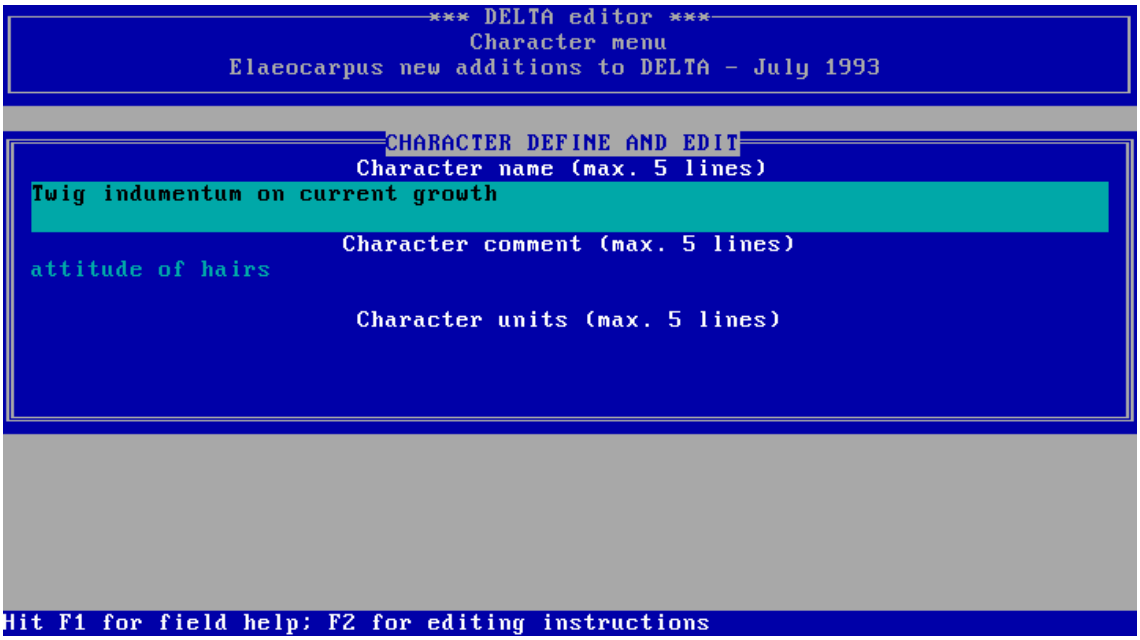
Screen 1: Choosing a DELTA data file



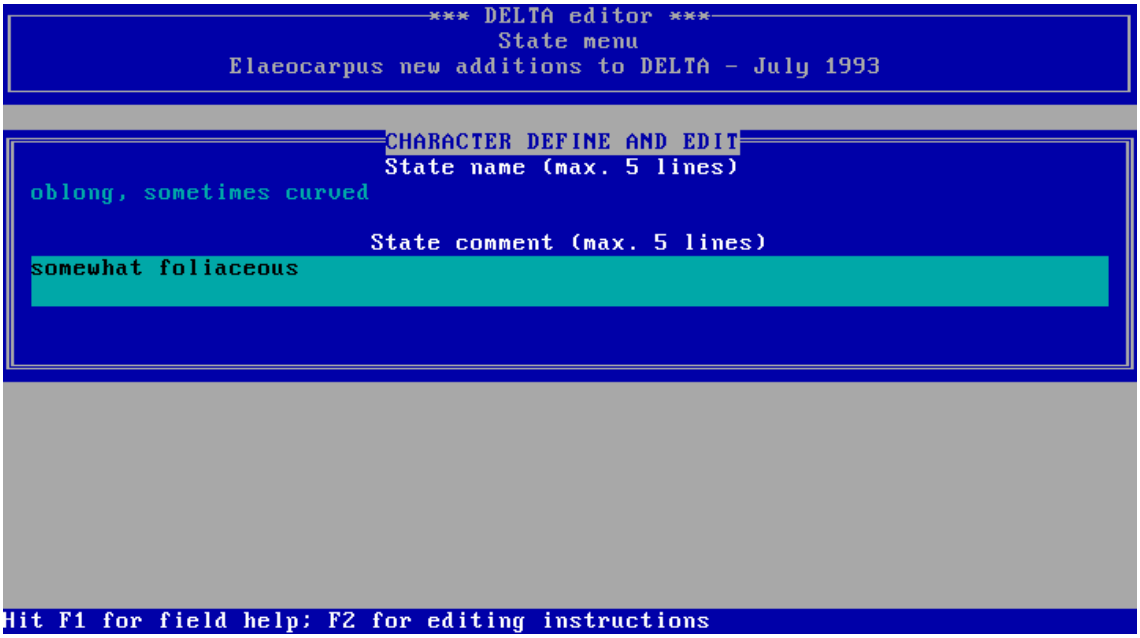
Screen 2: Selecting a character from a loaded DELTA set



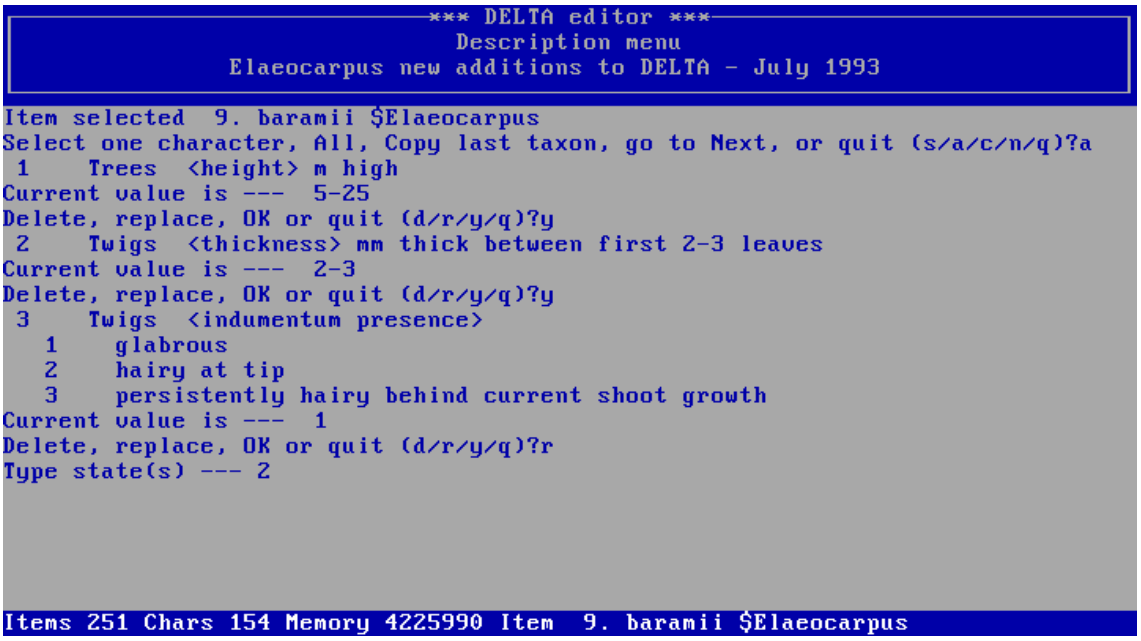
Screen 3: Editing the definition of a character



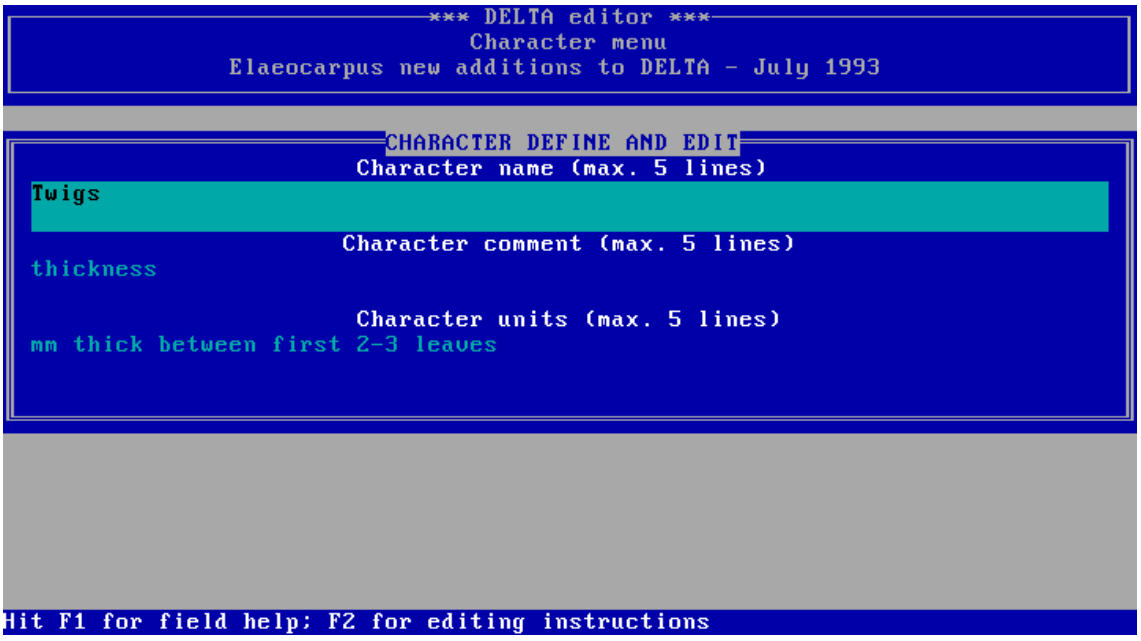
Screen 4: Editing a character state



Screen 5: Making changes to the description of a taxon



Screen 6: Adding the defintion of a new character



General references

PANKHURST, R.J. (1991) “Practical taxonomic computing”. Cambridge University Press, 202pp.

PANKHURST, R.J. (1991) Algorithms for identification. In “Symbolic-Numeric Data Analysis & Learning” eds. E. Diday & Y. Lechevallier, Conference INRIA, Versailles Sept. 1991, pp.3-13. Nova Science Publishers Inc., New York

PANKHURST, R.J. (1993) Principles and problems of identification. In “Advances in Computer Methods for Systematic Biology” ed. R.Fortuner, pp. 125-136. Johns Hopkins UP, Baltimore & London.

PANKHURST, R.J. (1998) A historical review of identification by computer, In "Information Technology, Plant Pathology & Biodiversity" eds. P.Bridge, P.Jefferies, D.R.Morse & P.R.Scott, pp. 289-303. CAB Intl., Wallingford.

Appendix

DELTA data file for the standard PANKEY example, *Jurinea* (Compositae) in Europe.

```
*HEADING JURINEA TEST/
*KEY OPTIONS  POLYCLAVE=21/
*NUMBER OF CHARACTERS 24
*PRINT WIDTH 75
*MAXIMUM NUMBER OF STATES 7
*MAXIMUM NUMBER OF ITEMS 17
*CHARACTER TYPES 2,RN 3,OM 14,IN 15,RN 20,RN 22,OM 23,OM 24,RN
*NUMBERS OF STATES 3,3 7,6 9,3 10,3 14,0 16,4 17,3 19,7 21,4 22,3 23,3
*CHARACTER WEIGHTS 3,0 5,0
*KEY STATES 2,0-10/10-200 14,1-4/5-100 15,0-3/3-100
            20,0-2.0/2.0-5.0/5.0-100  24,0-1.0/1.-2./2.-4./4.-100
*CHARACTER DESCRIPTIONS
#1. Stem <presence>/
 1. absent/
 2. present/

#2. Stem <height>/ cm/

#3. Stem <leaf distribution>/
 1. leafless/
 2. leafy at base/
 3. leafy throughout/

#4. Stem <shrubbiness>/
 1. herbaceous/
 2. woody at base/

#5. Rhizome <presence>/
 1. absent/
 2. present/

#6. Sterile rosettes <presence>/
 1. absent/
 2. present/

#7. Basal leaves <shape>/
 1. linear/
 2. linear-lanceolate/
 3. lanceolate/
 4. spathulate/
 5. ovate/
 6. ovate-oblong/

#8. Basal leaves <cut>/
 1. entire/
 2. pinnatifid/

#9. Upper surface of leaves <colour>/
 1. green/
 2. white/
 3. grey/

#10. Basal leaves <hair above>/
 1. subglabrous above/
```

 2. arachnoid-tomentose above/
 3. setose above/

#11. Leaf margins <recurved>/
 1. plane/
 2. revolute/

#12. Cauline leaves <amplexicaul>/
 1. not amplexicaul/
 2. amplexicaul/

#13. Cauline leaves <auricles>/
 1. without auricles/
 2. auriculate/

#14. Capitula <no.>/

#15. Capitula <length>/ cm/

#16. Capitula <shape>/
 1. cylindrical/
 2. subglobose/
 3. hemispherical/
 4. obconical/

#17. Outer involucral bracts <habit>/
 1. erect/
 2. patent/
 3. recurved/

#18. Outer involucral bracts <texture>/
 1. herbaceous/
 2. coriaceous/

#19. Distal part of bracts <colour>/
 1. whitish/
 2. green/
 3. yellowish/
 4. brown/
 5. pink/
 6. reddish/
 7. purple/

#20. Achene <size>/ mm/

#21. Achenes <shape>/
 1. subcylindrical/
 2. tetragonal/
 3. obpyramidal/
 4. conical/

#22. Achene <hair>/
 1. glabrous/
 2. puberulent/
 3. hairy/

#23. Corona of achene <size>/
 1. absent/

```
 2. inconspicuous/
 3. conspicuous/

#24. Pappus <relative length>/ times achene/

*DEPENDENT CHARACTERS  1,1:2-4:12-13
*ITEM DESCRIPTIONS

#1.J.linearifolia/
 1,2 2,12-40 3,3 4,2 5,1 6,1 7,2/3 8,1 10,2 11,V 12,1 13,1 15,0.5-1.8
 16,1 17,1 18,1 19,5/6/7 20,3.5-4.5 21,2 22,1 23,3 24,2.5-3.5

#2.J.stoechadifolia/
 1,2 2,10-40 3,3 4,1 5,1 6,1 7,1 8,1 10,3 11,2 12,1 13,1 14,3-20 15,2.7-3
 16,1/4 17,1 18,1 19,1/5 20,3.5-4.5 21,2/4 22,1 23,3 24,1.5-2.0

#3.J.tzar-ferdinandii/
 1,2 2,15-30 3,3 5,2 6,V 8,1 9,1 10,3 11,2 12,1 13,1 14,5-8 15,1.5-2 16,4
 17,1 19,1/2 20,3-4 21,2 22,1 23,3 24,3-4

#4.J.pinnata/
 1,2 2,4-13 3,1/2 4,2 5,1 6,V 8,2 9,2/3 10,2 11,1 12,1 13,1 14,2-3
 15,1.5-2.3 16,4 17,1 19,6/7 20,3-4.5 21,2 22,1 23,3 24,3-4

#5.J.tanaitica/
 1,2 2,15-60 3,3 4,1 5,1 6,1 8,2 9,2/3 10,2 11,1 12,1 13,2 14,2-20
 15,1.2-1.5 16,4 17,1 18,2 19,3/4/7 20,3.5-4.5 22,1 23,2 24,2-2.5

#6.J.albicaulis/
 1,2 2,30-75 3,3 4,2 5,1 6,1 7,1/2 8,V 9,1/2 10,2 11,1 12,1 13,1
 15,1.8-2.5 16,1/2 17,1 18,V 20,6-7 22,1 23,2 24,1.5

#7.J.kirghisorum/
 1,2 2,10-25 3,2 4,2 5,1 6,2 7,3/4 8,1 9,2/3 10,2 11,2 12,1 13,1 14,1-3
 15,1.1-1.3 16,3 17,1 18,1 20,4-5 23,2 24,1.2-1.4

#8.J.cyanoides/
 1,2 2,20-70 3,3 4,1 5,1 6,1 8,V 9,1 10,1 11,1 12,1 13,1 14,1-3
 15,1-3 16,2 17,1 18,1 19,2 20,3-4 21,3 22,1 23,2 24,2-2.5

#9.J.ewersmanii/
1,2 2,20-70 3,3 4,1 5,1 6,1 8,V 9,1 10,1 11,1 12,1 13,2 14,1-3 15,1-3 16,2
17,3 18,1 19,2/7 20,5-6 21,1 22,1 23,2 24,2-2.5

#10.J.mollis/
 1,2 2,30-70 3,2/3 4,1 5,1 6,2 7,2/3 8,2 9,1/3 10,1/2 11,V 12,1 13,1
 14,1-8 15,2-6 16,2/3 17,2/3 18,V 19,7 20,3-5 22,1 23,3 24,1.1

#11.J.polyclonos/
 1,2 2,40-80 3,3 4,1 5,1 6,2 8,2 9,1 10,1 11,2 12,V 13,2 14,5-20 15,0.5-2
 16,2 17,2/3 18,1 19,7 23,2 24,0.8

#12.J.ledebourii/
 1,2 2,15-80 3,2 4,1 5,1 6,2 7,2 8,V 9,1 10,1 11,V 12,2 13,1 14,1-3 15,2-4
 16,2 17,2/3 18,1 20,1-2 22,2 23,2 24,1.1

#13.J.consanguinea/
 1,2 2,20-35 3,2 4,1 5,1 6,2 8,2 9,1 10,1 11,V 12,1 13,1 15,2-7 16,3 17,1
 18,V 22,1/3 23,2 24,1.1
```

```
#14.J.glycacantha/
 1,2 2,30-60 3,2/3 4,1 5,1 6,2 8,2 9,2/3 10,2 11,1 12,1 13,1 15,4-7.5 16,3
 17,2/3 20,1-2 21,4 24,1.2

#15.J.humilis/
 1,V 2,0-4 3,3 4,1 5,2 6,1 8,V 9,2/3 10,2 11,2 12,1 13,1 15,2-2.5 16,4
 17,2/3 18,1 19,2 20,3-7 22,3 23,2 24,5-7

#16.J.taygetea/
 1,V 2,0-4 3,3 4,1 5,2 6,1 8,2 9,2/3 10,2 11,2 12,1 13,1 15,2-2.5 16,4
 17,1 18,1 19,7 20,3-7 22,3

#17.J.fontqueri/
 1,V 2,0-4 3,1 4,1 5,2 6,2 7,5/6 9,1 10,1 11,1 12,1 13,1 15,3-4 16,2
 17,1 18,1 19,2 20,4-6 23,1 24,1.5-2

*END
```