

CAFCA

Chapter 1

Installation and Start-Up

© M. Zandee 1989,1996 All Rights Reserved.

February 1996

THIS PAGE INTENTIONALLY LEFT BLANK

I. INSTALLATION AND START-UP

A SHORT HISTORY OF CAFCA

CAFCA is an acronym that stands for **C**ollection of **A**PL **F**unctions for **C**ladistic **A**nalysis. As indicated in the acronym all algorithms implemented in CAFCA are written in APL. APL is developed by K.E. Iverson, originally as a notational tool for mathematics, but soon implemented on IBM mainframes as an interpreted computer language. Because APL code is interpreted, programs (or functions in APL jargon) written in APL can in terms of speed hardly compete with programs written in languages as C or Pascal that are, in general, implemented as compilers. This is especially the case for algorithms that contain loops. On the other hand, APL contains very powerful operators enabling the programmer to achieve in one line of code the equivalent of dozens of lines of C or Pascal.

CAFCA was originally developed for use on mainframe computers. These versions were never distributed. When APL interpreters for personal computers became available CAFCA was ported to PC in 1985. In 1986 a menu driven user-interface was programmed and CAFCA could be distributed, eventually, as an integrated APL workspace. The main obstacle in its distribution was the need for users to separately acquire the APL interpreter as well, in order to make CAFCA run. Only after an APL interpreter with a free runtime license became available for Atari and Macintosh computers, the constraints on CAFCA's distribution could be lifted and CAFCA was ported to the Atari (1988) and to the Macintosh (1989). Soon further development was focused on the Macintosh only, including a native version for the PowerPC chip (1994). A Windows version is planned for 1996. The conversion of the CAFCA kernel will be relatively easy but the development of a new user-interface will take some time.

LICENSE AGREEMENT

Copyright (c); 1988-1996 M. Zandee.

Permission to use and distribute this software and its documentation for educational and research purposes is hereby granted without fee, provided the above copyright notice, author statement and this permission notice appear in all copies of this software and related documentation.

THE SOFTWARE IS PROVIDED "AS-IS" AND WITHOUT WARRANTY OF ANY KIND, EXPRESS, IMPLIED OR OTHERWISE, INCLUDING WITHOUT LIMITATION, ANY WARRANTY OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

IN NO EVENT SHALL THE AUTHOR, THE INSTITUTE OF EVOLUTIONARY AND ECOLOGICAL SCIENCES, OR LEIDEN UNIVERSITY BE LIABLE FOR ANY SPECIAL, INCIDENTAL, INDIRECT OR CONSEQUENTIAL DAMAGES OF ANY KIND, OR ANY DAMAGES WHATSOEVER RESULTING FROM LOSS OF USE, DATA OR PROFITS, WHETHER OR NOT ADVISED OF THE POSSIBILITY OF DAMAGE, AND ON ANY THEORY OF LIABILITY, ARISING OUT OF OR IN CONNECTION WITH THE USE OR PERFORMANCE OF THIS SOFTWARE.

The Macintosh version of CAFCA was developed using APL.68000, a proprietary product of MicroAPL Ltd., which has given permission for a runtime version of APL.68000 to be distributed with the software, or to be packaged into the software as one executable. Copyright and all intellectual property rights in APL.68000 remain vested in MicroAPL Ltd.

APL.68000 is a trademark of MicroAPL Ltd. Apple and Macintosh are trademarks of Apple Computer, Inc.

DISTRIBUTION DISK

Depending on the distribution version of CAFCA, the number and type of files in your distribution copy may vary. If there is a `Read me first` file present, please read this file first for notification of last minute changes and possible modifications of the installation procedure.

A note on Mac models:

Some of the older 68030 models [SE/30, IICx, IICI] have an accompanying 68882 FPU, next to the 68030 CPU, some don't! Standard 68040 CPU's, as used in the Quadra's 700, 800, 840, 900, and 950, have them built-in. The 68LC040's used in some other Mac models [e.g., Quadra/Performa 605 and 630, some of the powerbooks] do not have a FPU! Check your Mac's manual. In emulation mode the PowerPC line of Macintoshes are considered to have no FPU [although they do!]. The native version of the APL interpreter recognizes the 601, 603, and 604 PowerPC CPU and its floating point unit..

System 6 users:

Your distribution copy contains 6 items:

- *. a file named `Read me first`. Please do what the name suggests.
- *. an file with the name `CAFCA`.
- *. Two folders, one called `Help`, and the other `Xmpls`. The first contains `Help` files; the latter contains examples.
- *. Two files, one named `APL.Run.nofp` and the other `APL.Run.fp`.

This is a basic CAFCA configuration that will run on virtually all types of Mac's (including Power PC's), either under system 6 or system 7.

Your copy contains two different versions of the APL Runtime interpreter. The one named **APL.Run.nofp** is meant for Mac's without a floating-point co-processor installed. The one named **APL.Run.fp** is meant for Mac's with a floating-point co-processor next (68881) to or built into (68040) their CPU's. After unpacking, trashcan the files you don't need. No disasters will happen if you do not use the appropriate runtime version, although the FP version won't run on a Mac without a FPU. The non-FPU version will run on any configuration, including Mac's that still use system 6. The point is that if your CPU has floating-point support, use the software that can see and use it when needed.

System 7 users, including Power Macintosh:

Your distribution copy contains 4 items.

- *. a file named `Read me first`. Please do what the name suggests.
- *. an executable file (program) with the name `CAFCA/mac` or `CAFCA/ppc` (plus a version indication).
- *. Two folders, one called `Help`, and the other `Xmpls`. The first contains `Help` files; the latter contains examples.

In the `Xmpls` folder, several `TEXT` files are present. They are identical with the example data used in the manual. They give you an idea of how CAFCA input should look like, and you can use them to run the examples.

ANIMAL.DST	A binary matrix representing the distribution of taxa (rows) from ANIMAL over areas (columns). To be used in a biogeographic analysis to derive an area-data matrix,
------------	--

	jointly with ANIMAL.TRE as representing the cladogram.
ANIMAL.ars	A binary matrix representing the distribution of areas (rows) from ANIMAL over taxa (columns). To be used in a biogeographic analysis to derive an area-data matrix, jointly with ANIMAL.TRE as representing the cladogram. The areas (rows) are provided with names. It eliminates the use of ANMLAREA.NMS
ANIMAL.TRE	A binary matrix representing a cladogram for 5 taxa (taxa in rows, nodes of cladogram in columns). Can be used in a biogeographic analysis, jointly with ANIMAL.DST, to derive an area-data matrix.
ANIMAREA.ASC	A text file with the multi-state expression of an area-data matrix for ANIMAL. It is equivalent to the use of the binary expression in conjunction with the partitioning vector ANIMAREA.CPV
ANIMAREA.BIN	Binary area-data matrix for ANIMAL. Can be used to run a biogeographic analysis. When used in conjunction with ANIMAREA.CPV a multi-state expression will result (Also available separately as ANIMAREA.ASC).
ANIMAREA.CPV	Partitioning vector for the binary area-data matrix ANIMAREA..BIN
ANMLAREA.NMS	A text file with the names of the areas (rows) in ANIMAL.DST
PLANT.ASC	Multi-state data matrix, with taxa represented by the rows and characters by the columns. To be used in a primary character analysis.
PLANT.NMS	Namelist for taxa in PLANT.ASC or PLANTB.INP
PLANTB.INP	Binary data matrix, with taxa represented by the rows, and character states by the columns. To be used for primary character analysis. The equivalent of PLANT.ASC except for an unresolved polytypism in character 9.
PLANTB.CPV	Column partitioning vector for PLANTB.INP. Without it CAFCA cannot know which columns in PLANTB.INP represent the character states within each character.
PLANT.ars	A binary matrix representing the distribution of areas (rows) from PLANTB.INP over taxa (columns). To be used in a biogeographic analysis to derive an area-data matrix, jointly with PLANT.TRE as representing the cladogram. The areas (rows) are provided with names. It eliminates the use of PLNTAREA.NMS
PLANT.DST	A binary matrix representing the distribution of taxa (rows) from PLANTB.INP over areas (columns). To be used in a biogeographic analysis to derive an area-data matrix, jointly with PLANT.TRE as representing the cladogram.
PLANT.TRE	A binary matrix representing a cladogram for 5 taxa (taxa in rows, nodes of cladogram in columns). Can be used in a user-tree evaluation, or in a biogeographic analysis.
PLNTAREA.ASC	A text file with the multi-state expression of an area-data matrix for PLANT. It is equivalent to the use of the binary expression in conjunction with the partitioning vector PLNTAREA.CPV

PLNTAREA.BIN	Binary area-data matrix for PLANT. Can be used to run a biogeographic analysis. When used in conjunction with PLNTAREA.CPV a multi-state expression will result (Also available separately as PLNTAREA.ASC).
PLNTAREA.CPV	Partitioning vector for the binary area-data matrix PLNTAREA..BIN
PLNTAREA.NMS	A text file with the names of the areas (rows) in PLANT.DST
SECOND.BIN	A binary data matrix, with many taxa relative to the number of characters. Its primary analysis will result in unresolved cladograms. Therefore it can be used in a secondary analysis (directly after its primary analysis, or, much better, after saving the results of the primary analysis to an OutputFile). All its columns are equivalent (characters with only one state) and it therefore needs no column partitioning vector.
SECOND.pau.trees	Cladograms for SECOND.BIN matrix as generated by PAUP.
Alcidae	Data matrix for the <i>Alcidae</i> (Aves) from Strauch (1984). Used in the comparison of CAFCA with character compatibility analysis.
Alcidae.nex	Idem, in NEXUS format. Ready to be used in PAUP
Meacham81	Data matrix from Meacham (1981; his table 1). Used in the comparison of CAFCA with character compatibility analysis.
Meach81.nex	Idem, in NEXUS format. Ready to be used in PAUP.
B&McLt716.dst	Distribution of 5 major groups of ostariophysans from table 7.16 in Brooks & McLennan (1991).
B&McLt716.tre	Their cladogram; fig 7.30 in idem.
B&McLt717a.dst	Distribution of tapeworm genus <i>Acanthobothrium</i> from table 7.17 in Brooks & McLennan (1991).
B&McLt717a.tre	Their cladogram; fig 7.32a in idem.
B&McLt717b.dst	Distribution of tapeworm genus <i>Potamotrygonocestus</i> from table 7.17 in Brooks & McLennan (1991).
B&McLt717b.tre	Their cladogram; fig 7.32b in idem.
B&McLt717c.dst	Distribution of tapeworm genus <i>Rhinebothroides</i> from table 7.17 in Brooks & McLennan (1991).
B&McLt717c.tre	Their cladogram; fig 7.32c in idem.
B&McLt727.dst	Distribution of amphilinid flatworms from table 7.27 in Brooks & McLennan (1991).
B&McLt727.tre	Their cladogram; fig 7.54 in idem.
B&McLt727.utr	The cladogram of their hosts; fig 7.56 in idem.
B&McLt728.dst	Distribution of the tapeworm genus <i>Alcataenia</i> from table 7.28 in Brooks & McLennan (1991).
B&McLt728.tre	Their cladogram; fig 7.57 in idem.
B&McLt728.host.tre	Independent cladogram for hosts (seabirds). To be used as user tree
B&McLt728.bin	Binary data matrix resulting from the combination of the tapeworm cladogram and the distribution of tapeworms over seabirds.
B&McLt728.cpv	Its partitioning vector for columns
B&McLt728.asc	Multi state expression of B&McLt728.bin

Brooks90f5.dst	Distribution of 8 species of amphilinid flatworms from figure 5 (table 4) in Brooks (1990).
Brooks90f5.tre	Their cladogram; fig 5 in idem.
Brooks90f7.dst	Distribution of 8 species of amphilinid flatworms from figure 7 (table 5) in Brooks (1990). (Indo-Malaysia represented twice).
Brooks90f7.tre	Their cladogram; fig 5 in idem.
Brooks90f13.dst	Distribution of 4 hypothetical species from figure 13 (table 8) in Brooks (1990).
Brooks90f13.tre	Their cladogram; fig 5 in idem.
Brooks90f18a.dst	Distribution of 5 hypothetical species from figure 18a (table 11) in Brooks (1990).
Brooks90f18a.tre	Their cladogram; fig 18a in idem.
Brooks90f18b.dst	Distribution of 4 hypothetical species from figure 18b (table 11) in Brooks (1990).
Brooks90f18b.tre	Their cladogram; fig 18b in idem.
Brooks90t11.bin	Binary area-data matrix based on the distributional data and cladograms of two taxa in fig 18a and 18b.
Brooks90t11.nex	Idem in NEXUS format
Brooks90t11.og.nex	Idem, but including a full-zero out-area
Brooks90t11.cpv	Partitioning vector for columns in the binary area-data matrix, Brooks90t11.bin.
Brooks90f21a.dst	Distribution over 6 hypothetical areas of 6 hypothetical species from figure 21a (table 13) in Brooks (1990).
Brooks90f21a.tre	Their cladogram; fig 21a in idem.
Brooks90f21b.dst	Distribution over 6 hypothetical areas of 7 hypothetical species from figure 21b (table 11) in Brooks (1990).
Brooks90f21b.tre	Their cladogram; fig 21b in idem.
Brooks90f24a.dst	Neotropical distribution of 7 species from the genus <i>Selenidera</i> from figure 24a (table 15) in Brooks (1990).
Brooks90f24a.tre	Their cladogram; fig 21a in idem.
Brooks90f24b.dst	Neotropical distribution of 7 species from the genus <i>Pionopsitta</i> from figure 24b (table 15) in Brooks (1990).
Brooks90f24b.tre	Their cladogram; fig 21b in idem.
Brooks90f24c.dst	Neotropical distribution of 3 species from the <i>Pteroglossus viridis</i> group from figure 24c (table 15) in Brooks (1990).
Brooks90f24c.tre	Their cladogram; fig 21c in idem.
Brooks90f24d.dst	Neotropical distribution of 6 species from the genus <i>Pteroglossus bitorquatus</i> group from figure 24d (table 15) in Brooks (1990).
Brooks90f24d.tre	Their cladogram; fig 21d in idem.
Crusea.asc	Data matrix for <i>Crusea</i> from Estabrook & Anderson (1978).
Crusea.nex	Idem, in NEXUS format.
Gopher.tre	Cladogram for gophers as used in Page 1993, fig 11
Lice.dst	Distribution of chewing lice over species of gopher; Page 1993, fig 11.
Lice.tre	Cladogram for chewing lice; Page 1993, fig 11.
Page90.fig3.dst	Distribution of three hypothetical species over 4 hypothetical areas; from Page 1990, fig 3.
Page90.fig3.tre	The species cladogram; fig 3 in idem.

Page93.fig10.dst	Distribution of five hypothetical species over four hypothetical areas; from Page 1993, fig 10.
Page93.fig10.tre	The species cladogram; fig 10a in idem.
Page93.fig10.bin	The binary area-data matrix derived from the cladogram and distribution matrix in Page93.fig10.tre and Page93.fig10.dst
Page93.fig10.nex	Idem, in NEXUS format.

INSTALLATION

If you received your copy of CAFCA on a floppy disk, make a copy of your distribution disk. Store the original and use the copy.

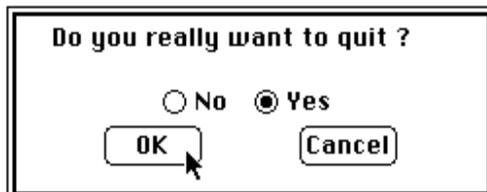
1. If there's a `Read me first` file present on the distribution disk, double click or select and open it for reading. If that file describes an installation routine, follow that recipe and skip the points 2 - 4 below.
2. Make a new folder on your hard disk and name it CAFCA.
3. Double click the self-extracting archived file with the name `CAFCA.mac7.sea` or `CAFCA.ppc.sea`, depending on the copy you received or downloaded. When prompted by the unpacking program, indicate the CAFCA folder you just made in step 2 to copy the extracted files onto.
That's all. You can skip step 4.
4. In case you used the `CAFCA.mac6.sea` archive you can delete one of the APL runtime interpreters, depending on the type of Mac you use (see the paragraph on distribution disk for system 6 users).

START-UP

Start your Mac with your own System™ and Finder™. To start CAFCA, double click the file name, or select the name and **Open** the file from the **File** menu. Once CAFCA is up you will see this opening screen, except that the **About CAFCA** window will contain your registration, unless you downloaded CAFCA from [//wwwbio.leidenuniv.nl/~zandee/cafca.html](http://wwwbio.leidenuniv.nl/~zandee/cafca.html).



Click anywhere on the screen to make the start-up window disappear. Now you are ready to run. If you want to **Quit** for now, select **Quit** from the **OutputFile** menu, or type Command-Q and confirm (**Yes**) in the next dialog.



THIS PAGE INTENTIONALLY LEFT BLANK