

Chapter 2

Working with trees

This chapter gives you information on basic tree operations. The topics covered include:

- reading and writing files of trees
- editing trees
- displaying and printing trees
- selecting subsets of trees
- pruning trees
- rooting trees
- computing tree statistics

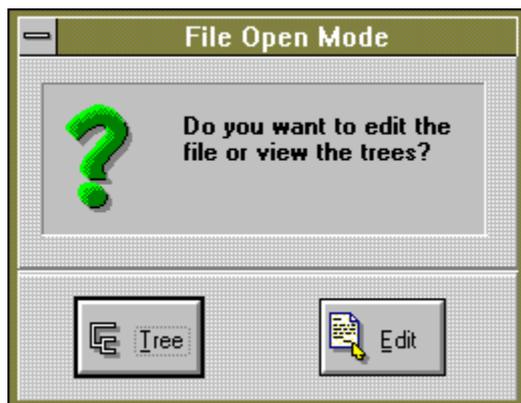
Tree input and output

This section describes how to input trees into COMPONENT from a disk file, and how to create a new file of trees. COMPONENT uses the NEXUS file format developed by David Swofford (1990) and Wayne and David Maddison (1992) for their programs PAUP and MacClade. This format is described in Chapter 3. COMPONENT also understands the file formats used by the programs COMPONENT 1.5, CONTREE, Hennig86, and PHYLIP.

Input files

You can input trees into COMPONENT using the **File Open** command. When you choose this command COMPONENT will display a standard File Open dialog box asking for the name of the NEXUS file you want to open. After you have entered the file name COMPONENT will display another dialog box asking whether you want to view the trees in the file or edit the file:

Figure 2.1
The File Open Mode
dialog box



If you choose **T**ree the program will read the trees in the file and then display the trees in a Tree window. If it encounters an error in the file it will display the file in an Edit window with the location of the error highlighted and an error message displayed on the status bar. To continue reading the file first correct the error then choose the **F**ile **E**xecute command.

If you choose **E**dit the program will display the file in an Edit window. You can subsequently input the trees into the program by choosing the **F**ile **E**xecute command.

In the input file you can explicitly describe the trees using TREE or UTREE commands, generate random trees with the ALL and RANDOM commands, or you can tell COMPONENT to import the trees from another file by using the IMPORT command.



*You can also generate trees from within COMPONENT without having to get them from a disk file (for example by using the **R**andom trees and **A**ll trees commands from the **G**enerate menu—see Chapter 6).*

Information on the trees

A NEXUS file may contain information about its contents in the form of "output comments" (see Chapter 3). You can view output comments using the **A**bout command from the **T**rees menu.

Reading NEXUS files from other programs

At present COMPONENT does not understand the DATA block used by PAUP and MacClade to store character data, hence no information in that block (such as taxon labels) can be extracted by COMPONENT. However, both PAUP and MacClade can write NEXUS files that COMPONENT can successfully read if you take the following precautions.

PAUP NEXUS files

PAUP tree files should be saved without a translation table (having a translation table assumes the existence of a TAXA or a DATA block). These files can be read by COMPONENT. If you wish to read a NEXUS file that has a translation table but no TAXA block you must use the IMPORT command with FILETYPE=NEXUS.

MacClade NEXUS files

MacClade NEXUS files should be saved with TAXA and CHARACTERS blocks instead of the DATA block. COMPONENT extracts what it needs from the TAXA block then skips the CHARACTERS block.



COMPONENT requires that each tree description contains the same number of taxa. Unlike PAUP, it does not join any missing taxa to the base of the tree, nor does it allow for unequal sized trees as does MacClade. These differences may be important if you are reading files originally intended for either PAUP or MacClade.

Importing trees from other programs

COMPONENT can read files written by a number of other programs, including Hennig86 and PHYLIP. This enables you to read in trees produced by those programs with a minimum of effort. Simply create an input file consisting of a TREES block containing just the IMPORT command. For example, this file will read in the Hennig86 tree file \HENNIG\MY.TRE..

```
#NEXUS

BEGIN TREES ;
    IMPORT FILETYPE=HENNIG86
           FILE=' \hennig\my.tre ' ;
ENDBLOCK ;
```

For a detailed discussion of the various formats supported by COMPONENT and how the program interprets those formats see Appendix A.

Saving the entire file

If you choose the **Save** command from the **File** menu the entire file will be saved to disk, including the contents of the DISTRIBUTION and TREES blocks.

COMPONENT saves the file in two formats: the NEXUS format and a special binary format called a CPS file. CPS files can only be read by COMPONENT, and are designed to speed up the input of trees into the program.

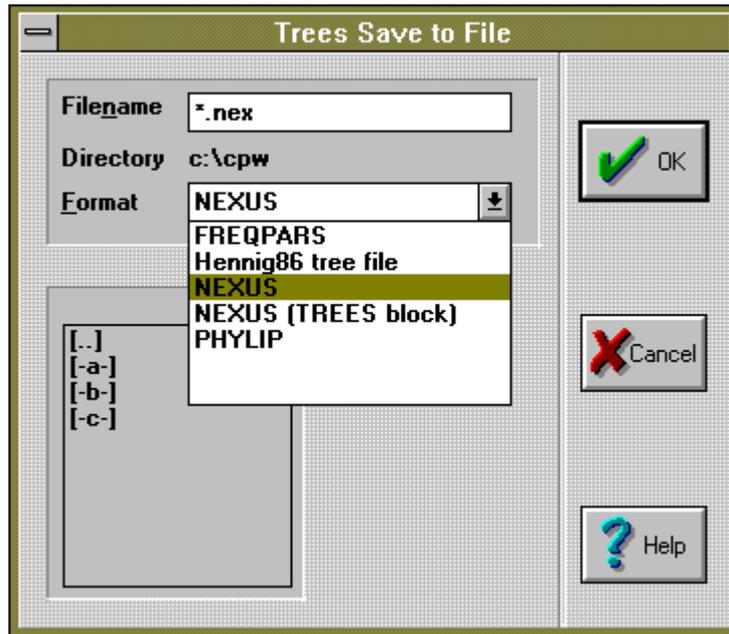


*COMPONENT reads CPS files much quicker than NEXUS files, so if your data file contains many trees using the **File Save** command to create a CPS file can substantially speed up subsequent input of those trees. The CPS file also saves the current leaf and tree activities.*

Saving trees in the current block

The trees in the current tree block can be saved using the **Save to file** command from the **Trees** menu. COMPONENT can write tree files that can be read by other programs, including FREQPARS, Hennig86, and PHYLIP, as well as standard NEXUS format files. When you chose the **Save** command a dialog box appears asking you for the name of the file and its format:

Figure 2.2
The Trees Save to File
dialog box



The default format is a NEXUS file with a TAXA and a TREES block. You can also write a NEXUS file containing just a TREES block. The other file formats presently supported are FREQPARS, Hennig86, and PHYLIP.



The programs FREQPARS and PHYLIP require binary (fully resolved) trees. If the set of trees you want to save contain nonbinary trees then these formats will not be available. For further details on the program formats supported by COMPONENT see Appendix A.

Editing trees

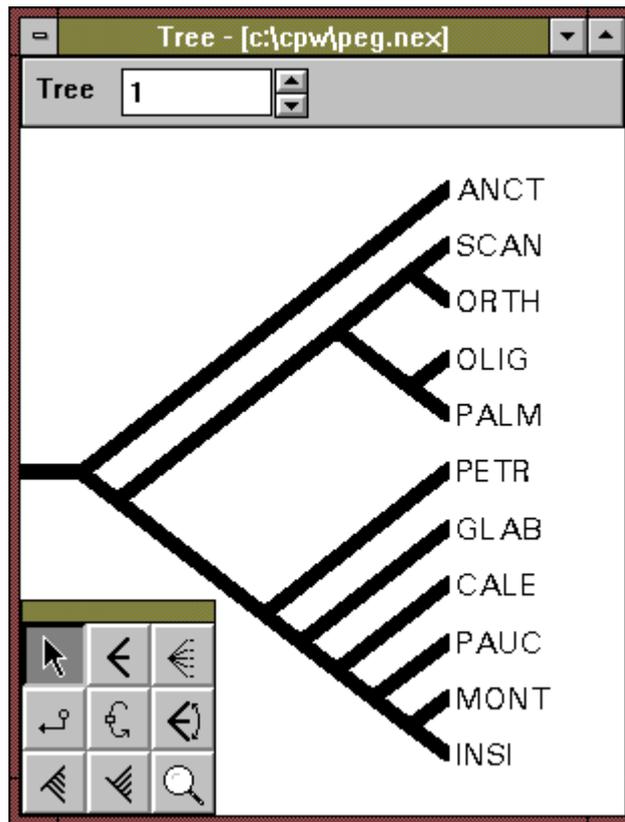
The trees stored in memory can be edited using COMPONENT's tree editor, which is modelled on the tree editor in MacClade 3.0 (Maddison and Maddison, 1992).

You can use the editor to change the topology of any tree in memory. For example, if you want to enter a tree from the literature, instead of typing in the complete tree you can create an input file with just the taxon names and an unresolved star tree, then use the editor to create the desired tree.

Opening the tree editor

To edit the tree displayed in the tree window choose the **T**ree command from the **E**dit menu. The tree window will display the tree with thick branches and a Tools palette showing you the tools available to manipulate the tree:

Figure 2.3
The Tree editor



The tree being edited is the original tree in memory. Any rooting or pruning operations you may have applied to the set of trees are ignored.

Closing the tree editor

While you are in edit mode you cannot view any other trees in the tree window, nor can you access the **Trees** menu until you close the editor. To close the editor choose **Tree** from the **Edit** command again. If you have modified the tree you will be asked whether you want to replace the original tree with the one you have created with the editor.

Tree manipulation tools

COMPONENT provides tools for moving branches, rerooting, polytomy formation, and rearranging the appearance of the tree. The tools available are displayed in the Tools palette:



Figure 2.4
The Tools palette

To select a tool click on the icon corresponding to the tool you want to use. Once you have selected a tool the cursor will change to the appropriate shape. To

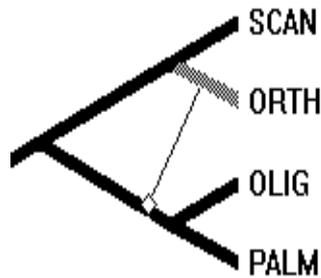
manipulate the tree click the cursor on a branch; the corresponding operation will be performed on that branch (unless the desired operation is invalid, in which case COMPONENT gently "beeps" and ignores the command). The default tool is the arrow tool, which is used for moving branches.



Move branch

A branch can be picked up and moved to another part of the tree by clicking on the branch and dragging it to the new position. The branch being moved will be highlighted with a different colour, and the cursor will change from an arrow to a small diamond:

Figure 2.5
Dragging a branch from
one location to another



Drag the diamond cursor to the position you want to move the branch, and release the mouse button. If the new placement is invalid (for example, moving a branch onto one of its descendants) COMPONENT will gently "beep" and ignore your request.



Collapse branch

Clicking with this tool on an internal branch deletes that branch produces a polytomy. COMPONENT ignores attempts to delete terminal branches.



Collapse clade

Clicking with this tool on an internal branch collapses the clade above that branch



Reroot

Clicking on a branch with this tool reroots the tree on that branch.

The next five tools change the appearance of the tree in the window, but do not affect the cladistic topology. They cannot be undone.



Rotate branches

This tool rotates the descendants of a binary node, so that the left descendant is now the right descendant, and visa-versa. You can only use this tool on binary internal nodes. If you want to rearrange the descendants of a polytomy you must use the next tool.

**Exchange branches**

To rearrange the order of the descendants of a polytomy use this tool to interchange pairs of branches. Click on one descendant, drag the cursor onto the other descendant, and release the mouse button.

**Ladderize left**

This tool orders the subtree rooted at the branch clicked on so that the "heavier" nodes (those nodes with the largest clusters) appear leftmost (uppermost on the screen).

**Ladderize right**

This tool orders the subtree rooted at the branch clicked on so that the "heavier" nodes (those nodes with the largest clusters) appear rightmost (lowermost on the screen).

**Expand/shrink clade**

You can expand the tree to get a close up of a clade by clicking on the branch below the clade with the magnifying glass tool. To shrink an expanded tree click with the tool on the basal branch.

Undoing a tree manipulation

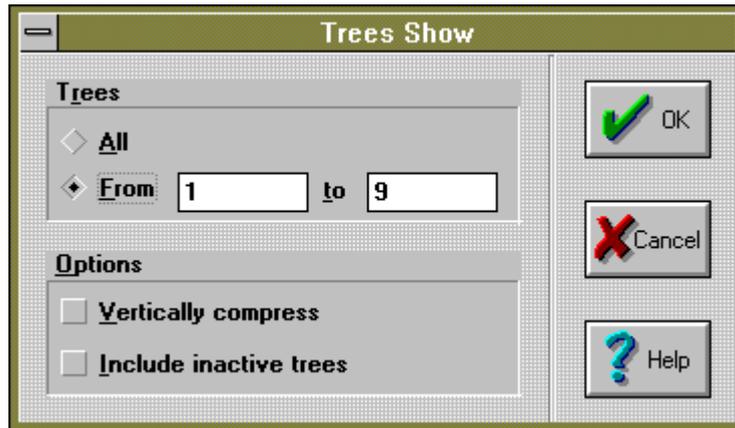
COMPONENT remembers the previous tree, so that you can undo a change and recover the old tree by choosing **Undo** from the Edit menu, or by pressing the *Alt + BackSpace* keys. Choosing the **Undo** command again restores the tree you just undid. **Undo** is only available for the operations that materially affect the tree topology: **move branch**, **collapse branch**, **collapse clade**, and **reroot**. Cosmetic changes (such as **rotate branches**) cannot be undone.

The **Undo** menu item will reflect which type of move can be undone; for example the menu item will read "Undo move branch" if you are using the **move branch** tool.

Displaying trees

The **Show** command from the **Trees** menu outputs to the display buffer the trees in the active Tree window. When you choose this command COMPONENT opens a dialog box enabling you to specify how many trees you want to show, and how they are to be drawn:

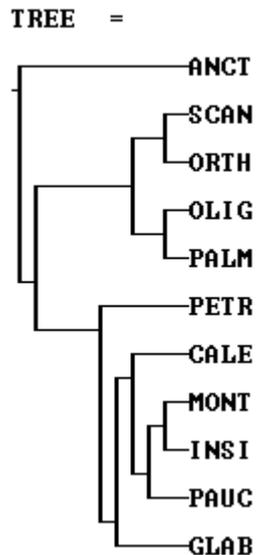
Figure 2.6
The Trees Show dialog box



By default, COMPONENT draws all the *active* trees in the current tree window. You can use the **From...to...** input boxes to specify a subset of trees you want displayed. To include inactive trees, check the **Include inactive trees** check box.

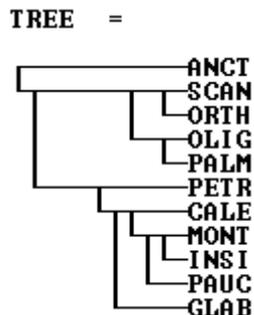
COMPONENT has two styles of drawing trees. The default is to have each node of the tree on a separate line. To conserve space on the buffer you can vertically compress the trees so that only the terminal nodes occupy separate lines (check the **Vertically compress** check box). Here is a tree drawn in the default style:

Figure 2.7
A tree drawn in the default style



Here is the same tree drawn vertically compressed:

Figure 2.8
The same trees as in
Figure 2.7 drawn
vertically compressed

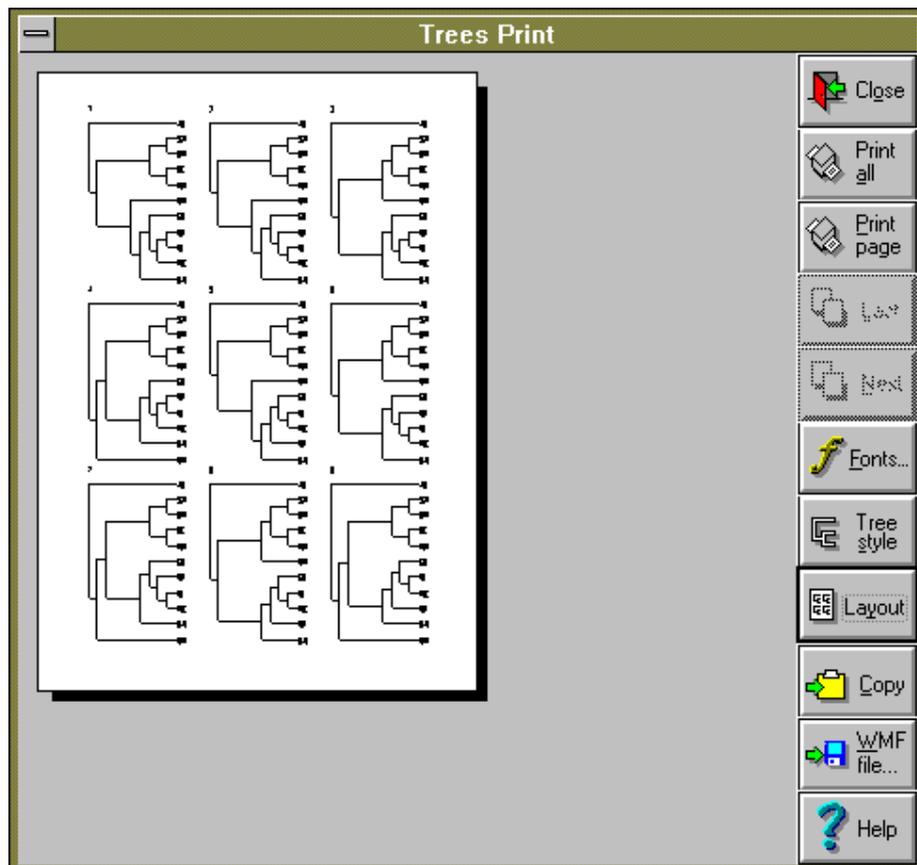


Printing trees

The trees stored in memory can be printed on the current system printer using the **Print** command from the **Trees** menu. You can also copy pictures of the trees to the Windows Clipboard, or save them to a graphics file.

Choosing the **Print** command causes COMPONENT to open a dialog box displaying a preview of how the trees will appear on the printed page:

Figure 2.9
The Trees Print dialog
box



The size and orientation of the page depend on the current settings of your printer. To alter these settings you should first use the **Printer setup** command from the **File** menu (see section "Printing" in Chapter 1).

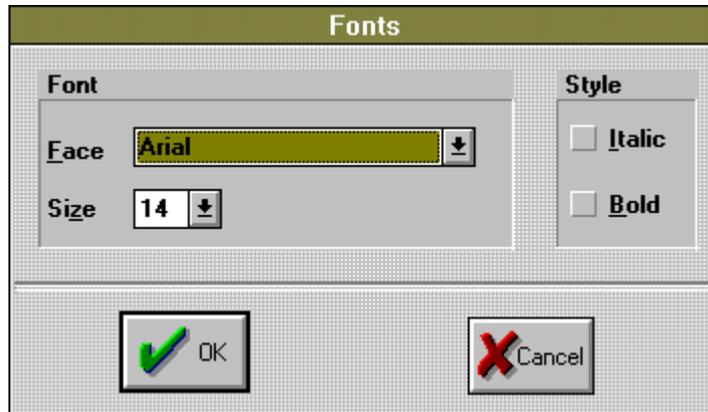
If there is more than one tree in the Tree window you can use the **Last** and **Next** buttons to move backwards and forwards through the pages. You can print either all the pages (**Print all**), or just the page being viewed (**Print page**).

The Trees Print dialog box provides a range of printing options: you can print single pages, place more than one tree on a page, choose between two styles of tree drawing, and adjust the font used to print the labels.

Changing the font

To change the font used for the labels click on the **Fonts** button. A dialog box will appear listing the fonts available for the current printer.

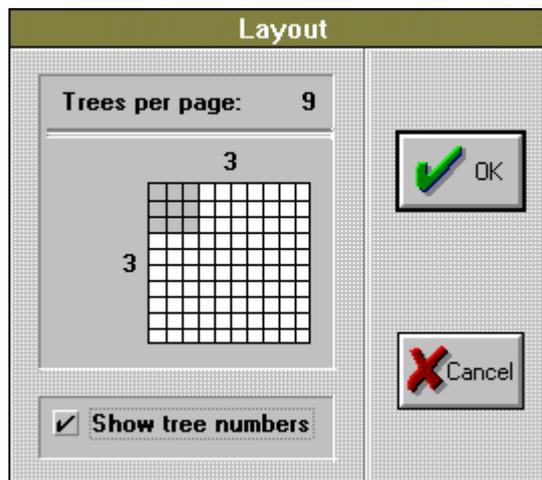
Figure 2.10
The Fonts dialog box



Printing more than one tree per page

Click on the **Layout** button. A dialog box will appear displaying a 10 ´ 10 grid representing the position of the tree pictures on the page. Click on the grid to set number and layout of trees on the page.

Figure 2.11
The Layout dialog box

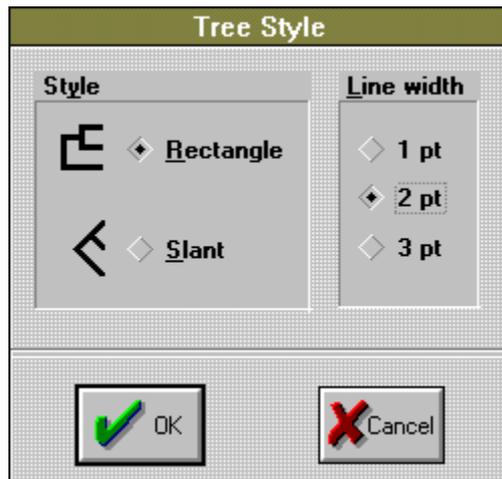


You can also toggle on and off the printing of tree numbers.

Changing the style of tree picture

COMPONENT supports two styles of tree picture, slanted branches and rectangular branches. To change the style of drawing click on the **Style** button. A dialog box will appear letting you choose between the two styles of picture. You can also choose three different line thickness.

Figure 2.12
The Tree Style dialog
box



Copying tree pictures to the Windows Clipboard

COMPONENT can paste the currently displayed page of trees to the Windows Clipboard as a *metafile picture*. You can paste this picture into any Windows application that recognises this format (for example Windows Write and Word for Windows). This makes it easy to include tree illustrations in publications.

To copy the page to the Clipboard click on the **Copy** button. The Clipboard now contains a picture of the trees. To place this picture into another application use that application's **Paste** command.

Saving tree pictures to a graphics file

COMPONENT can save the currently displayed page of trees as a WMF (Windows Metafile) graphics file. This file format is supported by most, if not all Windows graphics programs, as well as many word processors.

To create a graphics file click on the **WMF file** button. You will be prompted for the name of the file you wish to create.

Selecting trees

There may be times that you want to work with just a subset of the trees in memory. You can do this in COMPONENT by choosing the **Select** command from the **Trees** menu, which opens a dialog box:

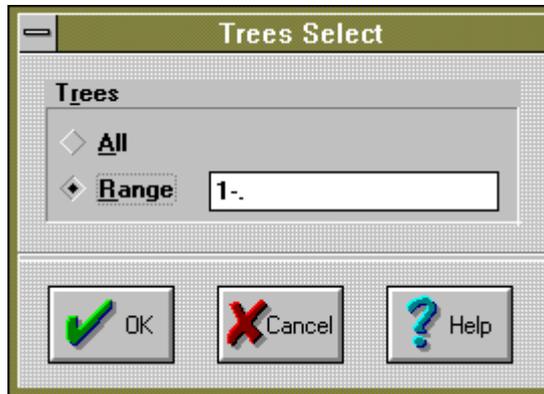


Figure 2.13
The Trees Select dialog
box

Selecting all trees

If you've previously selected a subset of trees you can reselect all the trees. Either

- click on the **All** radio button, or
- click on the **Range** radio button and type "all" in the input box.

Selecting a subset of trees

Click on the **Range** radio button to select a range of trees. Use the input box to enter your chosen range. You can list each tree separately, or specify a range of trees using hyphens to represent intermediate values. Examples of valid ranges are:

```
1 3
1-4
2 4-8 20-45
all
```

A period (.) represents the number of the last tree in memory, so that if there are 50 trees in memory the range

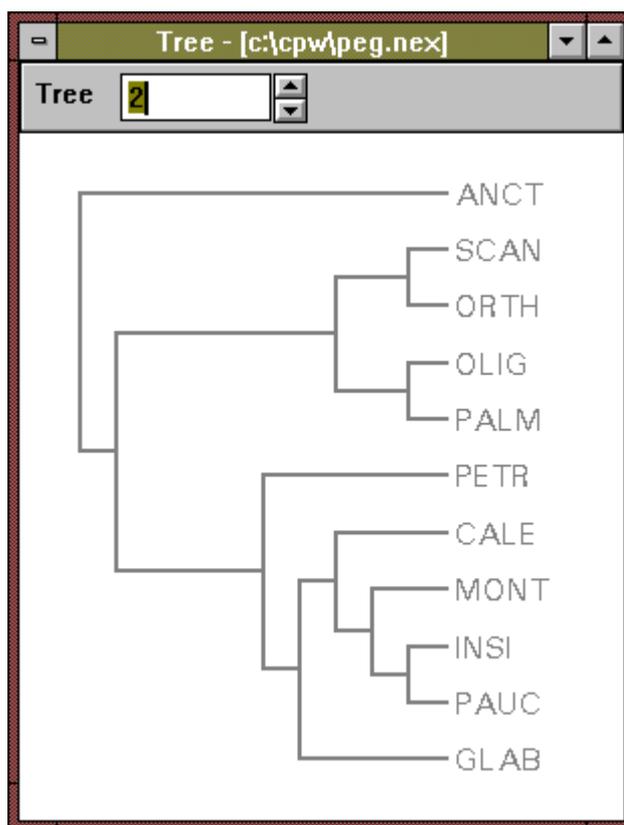
```
12- .
```

selects trees 12 through 50 inclusive.

Active trees

The trees currently selected are referred to as the *active* trees, the remaining trees are referred to as *inactive*. All subsequent **Trees** commands (with the exception of **Print**) work with just the active trees. Inactive trees are drawn in light grey in the Tree window:

Figure 2.14
A Tree window
displaying an inactive
tree



☞ The trees that are no longer active are still in memory, and can be accessed again by using the **Select** command again. Whenever a tree is referred to by number in COMPONENT's output that number refers to the tree's order in the complete set of trees.

Removing duplicate trees

The **Remove duplicate trees** command enables you to inactivate any tree that duplicates an already existing tree in the same block. This is useful, for example, if some of the trees differ solely in the placement of taxa which you have pruned using the **Prune/regraft leaves** command (see the next section). As with the **Trees Select** command, any trees inactivated are still in memory and can be made active again.

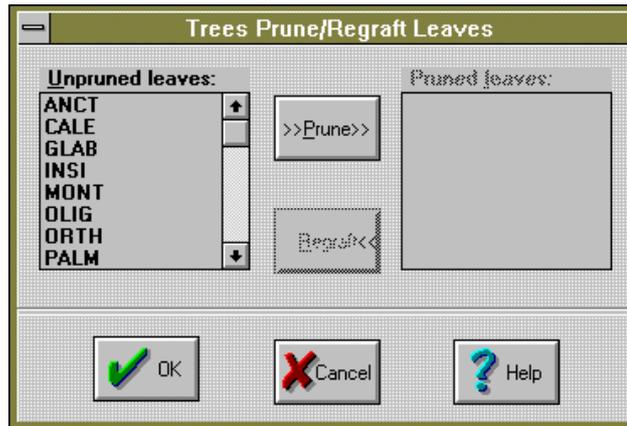
Pruning trees

COMPONENT allows you to alter the size of the trees by pruning leaves, or by grafting back on leaves previously pruned.

Pruning and regrafting leaves

Choose the **Prune or graft leaves** command from the Trees menu. COMPONENT will then display a dialog box with two scrollable lists. The current leaves are listed on the left, any leaves that have previously been pruned are listed on the right:

Figure 2.15
The Trees Prune/Regraft
dialog box



Using your mouse you can move leaves from one list to the other by either:

- Double-clicking on the desired leaf in one list box.
- Selecting a leaf label by clicking on it, then clicking on either the **Delete** or **Restore** button.



With the keyboard, use the *Tab* key to move to the list box you want, and use the \uparrow and \downarrow keys to select the leaf label you want to move. To move the label from one box to the other press *Alt+* the underlined letter of the desired button. For example, pressing *Alt+D* moves a label from the unpruned list box to the pruned list box.

COMPONENT checks whether the newly pruned (or regrafted) tree can still be rooted by the current outgroup (if any), and whether all the trees in the profile are binary or not. If the outgroup is no longer valid, COMPONENT issues a warning message and unroots the trees.



Pruning leaves does not affect the original trees which are still in memory, so you can restore the original trees at any time by unpruning the leaves. All tree output, consensus, and comparison commands use the pruned trees.

Rooting trees

Initially the trees in the Tree window are either rooted or unrooted, depending on how the trees were described in the input file (e.g., whether they were described with TREE or UTREE commands). The current rooting method is indicated by a check mark next to the appropriate item in the **Trees | Rooting** sub menu:

Figure 2.16
The Trees Rooting sub
menu

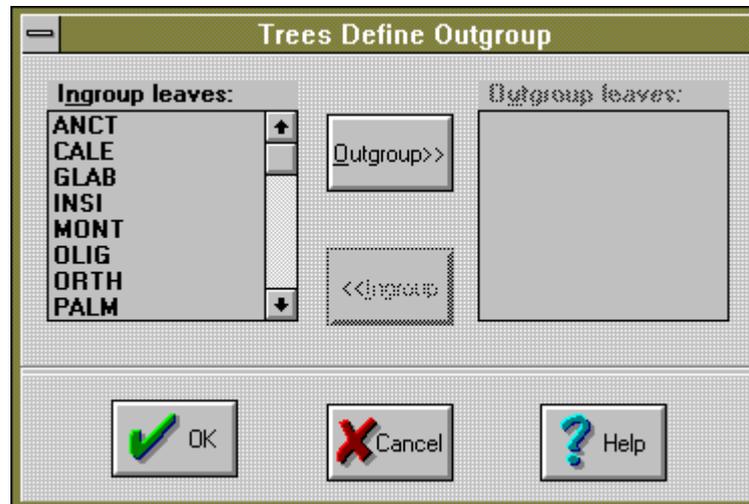


To choose a new rooting method, click on the desired menu item. If you have defined an outgroup you can also use that to root your trees.

Defining an outgroup

Choose the **Outgroup** command from the **Trees** menu. COMPONENT displays a dialog box with two scrollable list boxes. The current ingroup is listed on the left, the current outgroup on the right:

Figure 2.17
The Trees Define Outgroup dialog box



You can move leaves between the ingroup and outgroup list boxes in the same way as described for the Trees Prune/Regraft dialog box described above.

COMPONENT checks whether the outgroup you have chosen can be used to root all the trees in the profile. If it cannot, COMPONENT ignores the outgroup and unroots the trees in the profile.



*The **Outgroup** command only defines the outgroup. To use the outgroup to root the trees you must use the **Root with outgroup** command.*

Ordering trees

COMPONENT lets you change the order of the leaves in trees in the active tree window. If you choose the **Order** command from the **Trees** menu you will see the following dialog box:

Figure 2.18
The Trees Order dialog box



You have four kinds of ordering to choose from:

Input

Input ordering arranges a tree so that the order of the leaves in the tree agrees as closely as possible with their order in the input file.

Ladderize left

Trees ordered in this manner always have the largest subtree of each node as the left descendant of that node.

Ladderize right

As for ladderize left but with the largest subtree being the right descendant of its ancestor.

LLR

If the trees are all binary and have fewer than 15 leaves you can order the trees using Furnas' (1984) LLR method (see Appendix B). This method makes it easy to tell whether two trees have the same unlabelled topology ("shape") since their unlabelled trees will be identical (this need not be the case with the ladderizing options).

Tree statistics

At present COMPONENT computes three tree statistics, two of which apply solely to binary (fully resolved) trees: the *Sourdis number* and the *Topology number*. The Sourdis number is a unique identification number for a tree, the Topology number is a unique identification number for the unlabelled topology ("shape") of the tree. The final statistic applies to nonbinary trees and is the number of possible binary resolutions that are consistent with an incompletely resolved tree. To compute these statistics choose the **Statistics** command from the **Trees** menu.

There are several other tree statistics in the literature but I haven't incorporated them into COMPONENT as most confound tree information with tree shape (see Page, 1992).

Sourdis number

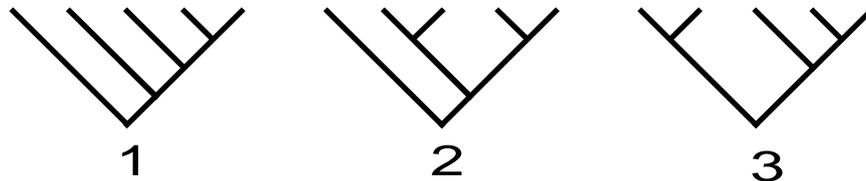
Sourdis (1985) described a method for numbering the set of possible rooted and unrooted binary trees that assigns a unique number to each tree (see also Rohlf, 1983). For example, each of the 15 possible rooted trees for four taxa would be assigned a number from 0 to 14. Sourdis developed this measure to help keep track of the large numbers of phylogenies produced during a simulation study (Sourdis and Krimbas, 1987).

COMPONENT can compute the Sourdis number for trees with up to 18 leaves (the limit is imposed by the maximum integer value the 80x87 coprocessor can compute.)

Topology number

Furnas (1984) described a method for assigning a unique number to the set of possible unlabelled rooted and unrooted binary (fully resolved) trees (see Appendix B). For example, for five taxa there are three possible rooted, unlabelled topologies and their number in Furnas' scheme:

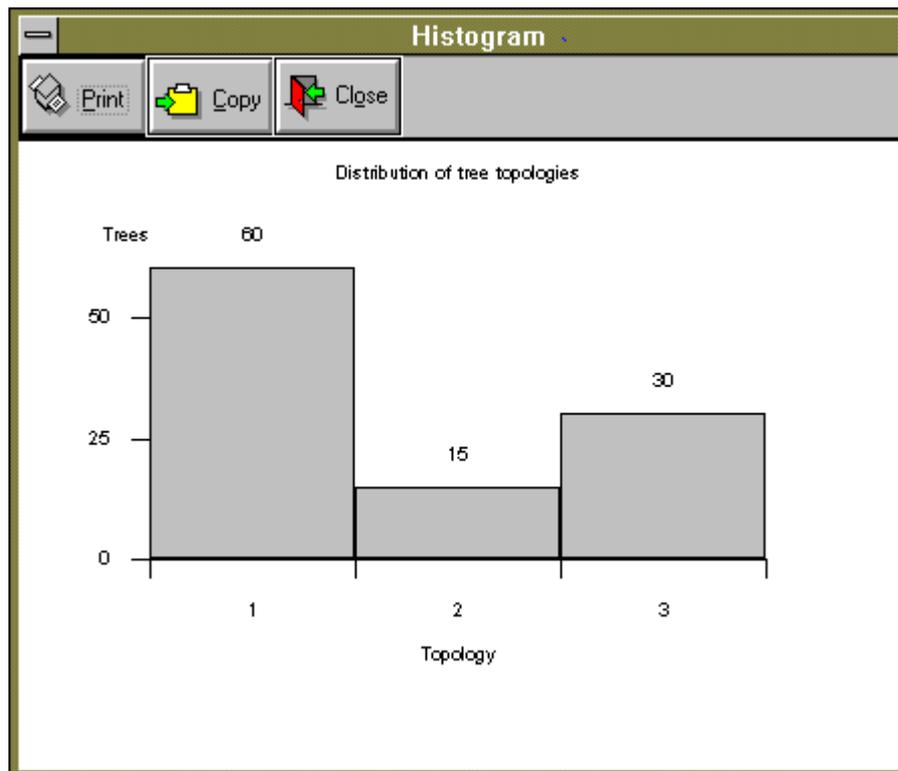
Figure 2.19
The three possible
rooted, unlabelled tree
topologies



For rooted trees Furnas (1984) numbering scheme corresponds to Harding's (1972) method for describing the shape of rooted binary trees. In Harding's notation the three trees above have shapes 5_1 , 5_2 , and 5_3 , respectively (see Page, 1993c).

Of the 105 labelled rooted binary trees for five taxa, 60 have topology 1, 15 have topology 2, and the remaining 30 have topology 3. COMPONENT will display the frequency of each topology in the set of currently active trees using a Histogram dialog box:

Figure 2.20
Histogram dialog box
showing frequency of
each of the tree shapes
in Figure 2.19 in the set
of 105 possible rooted,
binary trees for five taxa



Resolution

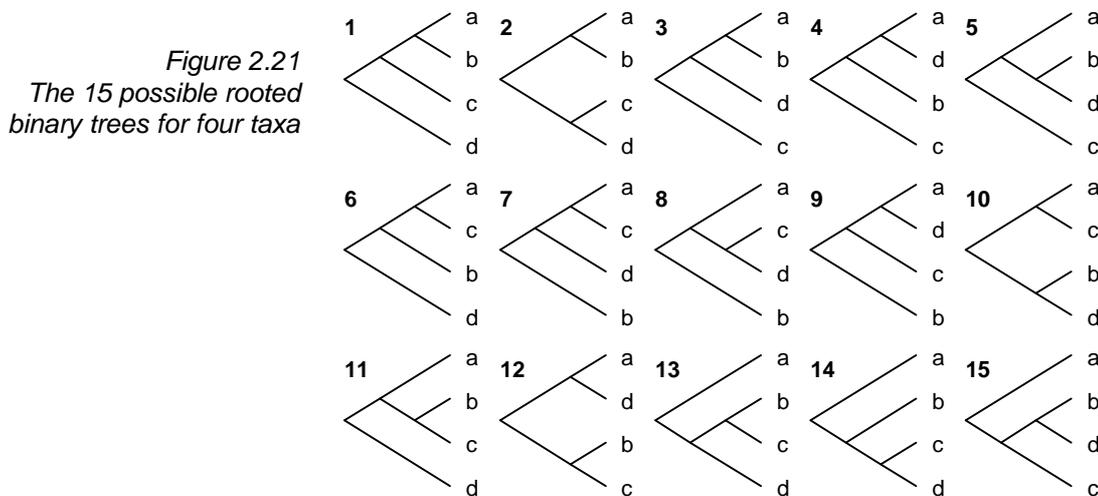
Rohlf (1982) presented a formula for computing the number of possible fully resolved trees that are consistent with a tree that contains one or more polytomies (normalised this value is his CI_2 index). For a binary tree there is just one possible resolution (itself), for a tree with one trichotomy there are three possible resolutions, and so on. A star tree (i.e., a completely unresolved tree or bush) has as many

resolutions as there are binary trees for that number of objects. Mickevich and Platnick (1989) refer to the complement of the number of possible resolutions as the tree's "prospective information" (see also Page, 1992).

COMPONENT will compute the number of allowed resolutions for trees with up to 18 leaves. If all the trees in the profile are binary then COMPONENT will not compute the number of resolutions (since they will all be equal to 1).

An example

The following example shows the output from an analysis of all 15 possible rooted trees for four taxa. Below are the trees generated using the **All trees** command from the **Generate** menu (see Chapter 6).



This is the output from the **Statistics** command:

Table 2.1
The Sourdis number and topology for the trees in Figure 2.21

Tree	Sourdis number	Topology number
1	14	1
2	12	2
3	13	1
4	10	1
5	11	1
6	3	1
7	4	1
8	2	1
9	0	1
10	1	2
11	8	1
12	5	2
13	9	1
14	7	1
15	6	1

Note that each tree has a unique Sourdis number from 0 to 14, and that 12 trees have shape **4**₁ and three trees have shape **4**₂.