Parallel Coordinate Plots towards Phylogenetic Trees

Rebecca Shapley
SIMS 247, Spring 2004
How to make a tree

Start with organisms.

To make an evolutionary tree, biologists take organisms, identify attributes that have different values due to evolutionary events, (characters and character states) and make a huge matrix encoding which state each organism has, then using algorithms to convert the matrix into a tree. My interest is in determining if parallel coordinates visualization can assist in the early part of the process, for identifying attributes/characters and values/states.
How to make a tree

Identify characters.

To make an evolutionary tree, biologists take organisms, identify attributes that have different values due to evolutionary events, (characters and character states) and make a huge matrix encoding which state each organism has, then using algorithms to convert the matrix into a tree. My interest is in determining if parallel coordinates visualization can assist in the early part of the process, for identifying attributes/characters and values/states.
How to make a tree

To make an evolutionary tree, biologists take organisms, identify attributes that have different values due to evolutionary events, (characters and character states) and make a huge matrix encoding which state each organism has, then using algorithms to convert the matrix into a tree. My interest is in determining if parallel coordinates visualization can assist in the early part of the process, for identifying attributes/characters and values/states.

How to make a tree

Apply algorithm to reconcile groupings and propose a tree.

To make an evolutionary tree, biologists take organisms, identify attributes that have different values due to evolutionary events, (characters and character states) and make a huge matrix encoding which state each organism has, then using algorithms to convert the matrix into a tree. My interest is in determining if parallel coordinates visualization can assist in the early part of the process, for identifying attributes/characters and values/states.
Data: 60 Salamanders

SVL: snout-vent length
TTL: total length
TAL: tail length
TAD: maximum tail depth
HL: head length
HW: head width
IO: distance between from corner of eyes
EN: distance from front corner of eye to nostril
IN: distance between nostrils
AX: distance between limbs on right side
AL: length of front limb
PL: length of back limb

-From Sean Rovito, IB, UC Berkeley
Phylogenetic Features

- Gaps
- Phylogenetically informative characters
- OTU signature
- Non-independence: combine for new one
- Non-independence: need to scale
Data: 60 Salamanders
Data: 60 Salamanders

Notice the Gap.
Data: 60 Salamanders

Notice the Gap.
Data: 60 Salamanders

Brushing for “A”, can see the a likely OTU and that A is probably a parsimoniously informative character.
Data: 60 Salamanders
Data: 60 Salamanders
Data: 60 Salamanders

Just the records with a positive slope between AX and AL are selected.
Data: 60 Salamanders

Notice - the lower values on most of the axes tend to be shared by the same lines…
As do the middle values… and the upper values. Lots of horizontal movement means a need to scale one value by other values to remove the redundant information about the absolute size of the organism, and focus instead on proportions.
Parallel Coordinates in Biology

Parallel Coordinates in Biology

The measurements across the bottom are standard body and skull measurements for small mammals.