

Not The Fly On The Wall:

can systematists cope with uncertainty?

Joe Felsenstein

Systematics Association: Huxley Lecture, 2008

What can we not hope to find out?

- Exact ancestral states (if no fossil molecules)

What can we not hope to find out?

- Exact ancestral states (if no fossil molecules)
- Exact past events in biogeography (if no relevant fossils)

What can we not hope to find out?

- Exact ancestral states (if no fossil molecules)
- Exact past events in biogeography (if no relevant fossils)
- Whether a single amino acid change was selected

What can we not hope to find out?

- Exact ancestral states (if no fossil molecules)
- Exact past events in biogeography (if no relevant fossils)
- Whether a single amino acid change was selected
- The precise sequence of horizontal gene transfer events

What can we not hope to find out?

- Exact ancestral states (if no fossil molecules)
- Exact past events in biogeography (if no relevant fossils)
- Whether a single amino acid change was selected
- The precise sequence of horizontal gene transfer events
- The exact sequence of coalescent events at one locus

What can we not hope to find out?

- Exact ancestral states (if no fossil molecules)
- Exact past events in biogeography (if no relevant fossils)
- Whether a single amino acid change was selected
- The precise sequence of horizontal gene transfer events
- The exact sequence of coalescent events at one locus
- The exact sequence of genome rearrangements leading to a chromosome

What can we not hope to find out?

- Exact ancestral states (if no fossil molecules)
- Exact past events in biogeography (if no relevant fossils)
- Whether a single amino acid change was selected
- The precise sequence of horizontal gene transfer events
- The exact sequence of coalescent events at one locus
- The exact sequence of genome rearrangements leading to a chromosome

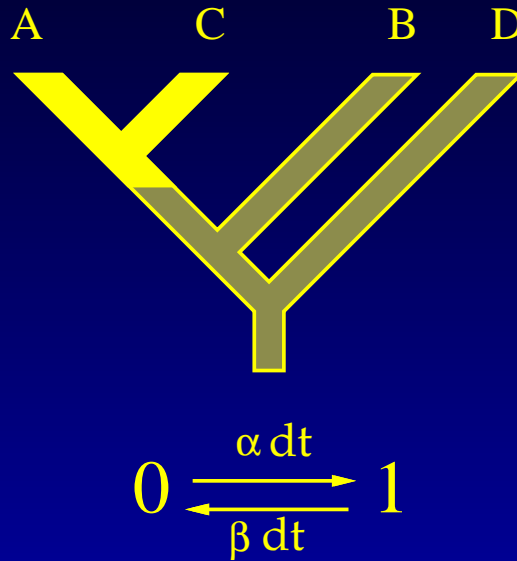
All these are “Fly On The Wall” questions about what actually happened, as opposed to questions about parameters of processes.

Can we describe uncertainty in morphology or behavior?

- Currently, most people using morphological or behavioral characters use parsimony methods
- They also code the characters discretely, which leads to the “character coding problem”
- They have no statistical model and little way of knowing how uncertain are their inferences (except for some bootstrapping or jackknifing which assumes characters change independently)
- Is discrete coding necessary? No.
- Are there statistical models for these discrete phenotypes? Yes.

Current methods for statistical treatment of 0/1 characters

Pagel (1994) and Lewis (2001) treat such data with



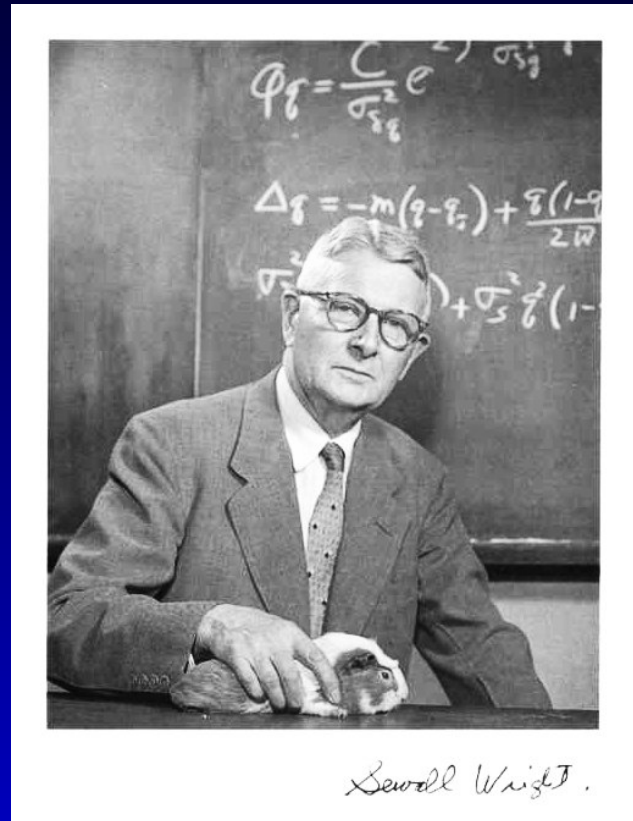
Pagel allows inference of whether change is correlated, on a known tree.

Lewis infers the tree, but does not allow for correlations among characters.

Neither takes into account contributions to a 0/1 character from multiple underlying loci.

The threshold model

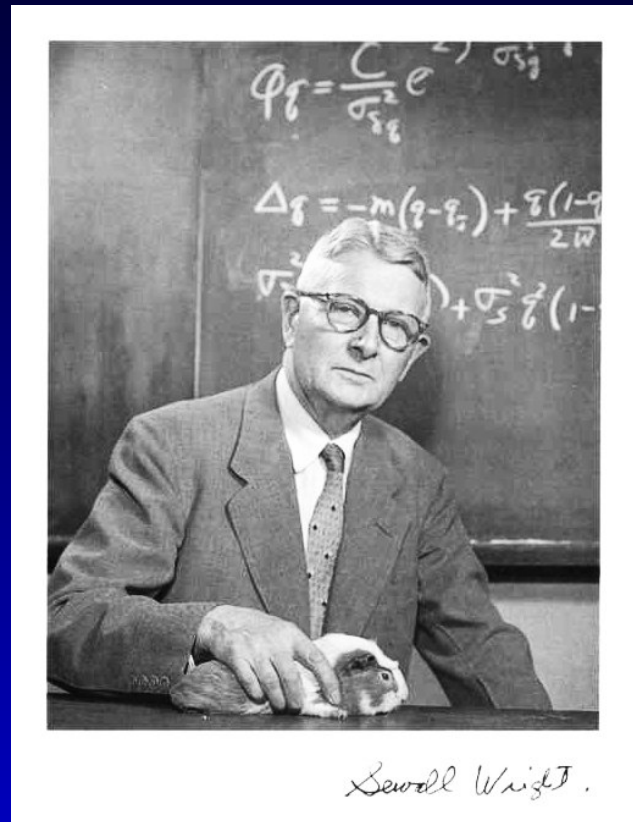
A relevant model was invented in 1934 by



Sewall Wright (1889-1988)
shown here in the late 1950's

The threshold model

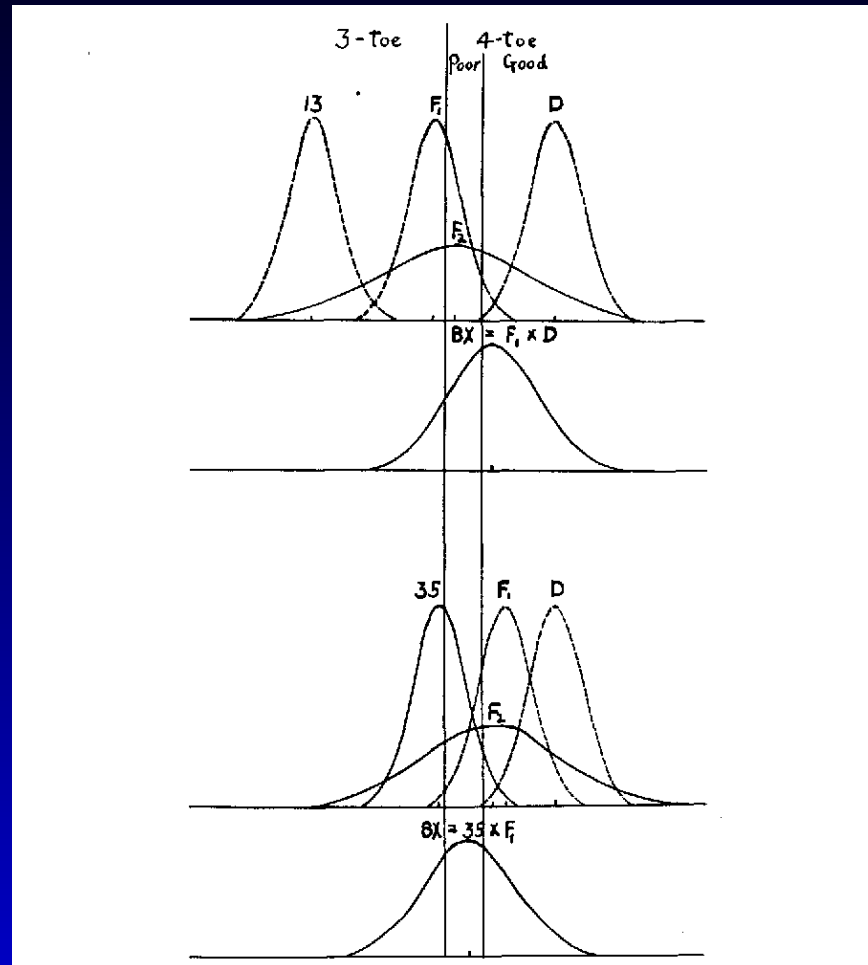
A relevant model was invented in 1934 by



Sewall Wright (1889-1988)
shown here in the late 1950's

(The story goes that he then absent-mindedly
started to erase the board with the guinea pig)

The threshold model, applied



Sewall Wright (1934), guinea pig digit number
(from Wright's follow-up 1934 second paper)

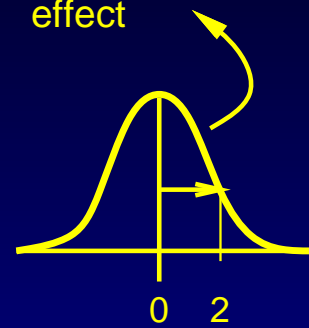
Brownian motion models for continuous characters

- Continuous characters change by natural selection, genetic drift, mutation and other population-genetic forces such as migration.
- The first three of these can be modelled by Brownian motion, with correlations among the changes in different characters.
- The correlations come from
 - Genetic covariation (which includes developmental and functional correlation, and the effects of pleiotropic mutations)
 - “Selective correlation” (Olaf Tedin, 1925; Stebbins, 1950) which is the covariation of selection pressures

The model is far from exact but is very tractable and a good starting point for attempts to make it more realistic.

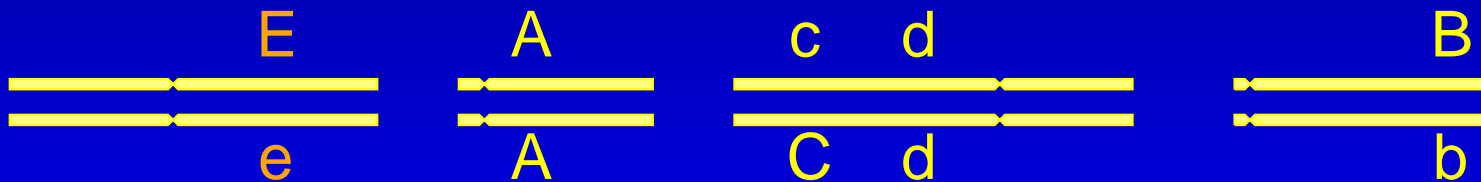
We use the standard quantitative genetic model

$$P = \left\{ \begin{array}{l} AA \ 2 \\ Aa \ 4 \\ aa \ 7 \end{array} \right\} + \left\{ \begin{array}{l} BB \ 0.6 \\ Bb \ 0.1 \\ bb \ -0.2 \end{array} \right\} + \left\{ \begin{array}{l} CC \ -1 \\ Cc \ 6 \\ cc \ 6 \end{array} \right\} + \left\{ \begin{array}{l} DD \ 0.3 \\ Dd \ 0.3 \\ dd \ 0.7 \end{array} \right\} + \left\{ \begin{array}{l} EE \ -0.4 \\ Ee \ 0.3 \\ ee \ -0.3 \end{array} \right\} + \text{environmental effect}$$



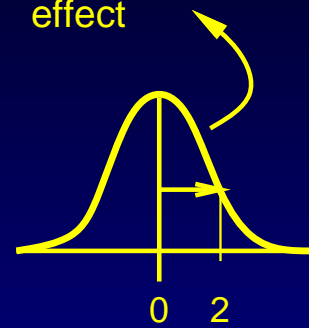
AA Bb Cc dd Ee
 Aa bb cc DD ee
 aa bb CC DD Ee
 aa bb Cc DD EE
 Aa Bb Cc DD Ee

$$0.3 + 2 + 6 + 0.7 + 0.1 + 0.9$$



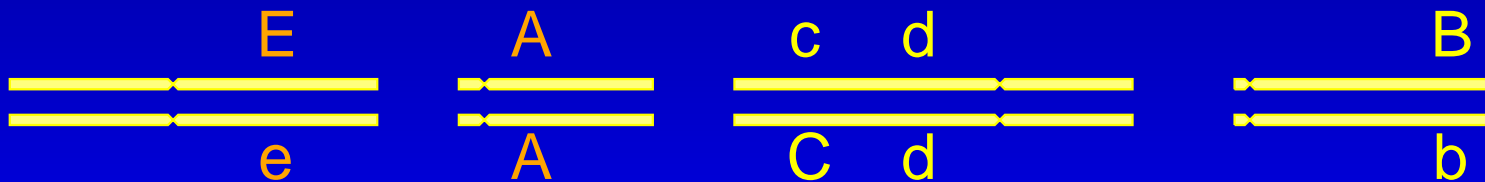
We use the standard quantitative genetic model

$$P = \left\{ \begin{array}{l} AA \quad 2 \\ Aa \quad 4 \\ aa \quad 7 \end{array} \right\} + \left\{ \begin{array}{l} BB \quad 0.6 \\ Bb \quad 0.1 \\ bb \quad -0.2 \end{array} \right\} + \left\{ \begin{array}{l} CC \quad -1 \\ Cc \quad 6 \\ cc \quad 6 \end{array} \right\} + \left\{ \begin{array}{l} DD \quad 0.3 \\ Dd \quad 0.3 \\ dd \quad 0.7 \end{array} \right\} + \left\{ \begin{array}{l} EE \quad -0.4 \\ Ee \quad 0.3 \\ ee \quad -0.3 \end{array} \right\} + \text{environmental effect}$$



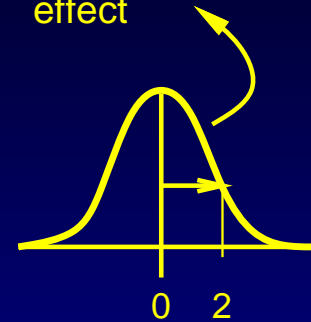
AA Bb Cc dd Ee
 Aa bb cc DD ee
 aa bb CC DD Ee
 aa bb Cc DD EE
 Aa Bb Cc DD Ee

$$0.3 + 2 + 6 + 0.7 + 0.1 + 0.9$$



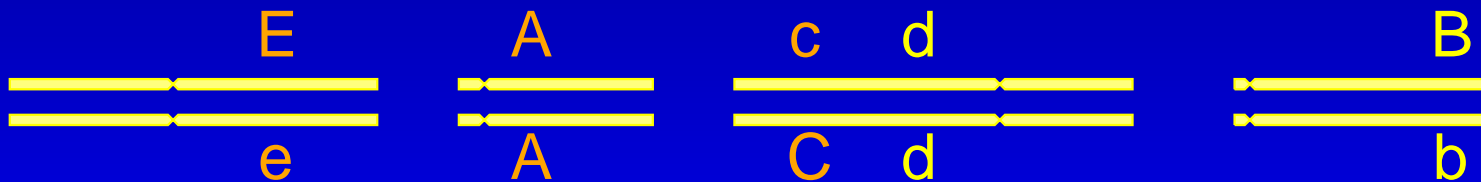
We use the standard quantitative genetic model

$$P = \begin{Bmatrix} AA & 2 \\ Aa & 4 \\ aa & 7 \end{Bmatrix} + \begin{Bmatrix} BB & 0.6 \\ Bb & 0.1 \\ bb & -0.2 \end{Bmatrix} + \begin{Bmatrix} CC & -1 \\ Cc & 6 \\ cc & 6 \end{Bmatrix} + \begin{Bmatrix} DD & 0.3 \\ Dd & 0.3 \\ dd & 0.7 \end{Bmatrix} + \begin{Bmatrix} EE & -0.4 \\ Ee & 0.3 \\ ee & -0.3 \end{Bmatrix} + \text{environmental effect}$$



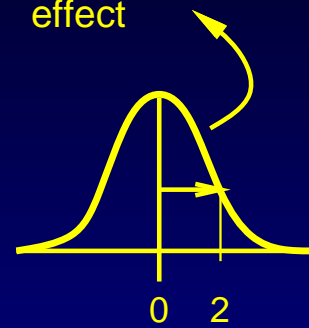
AA Bb Cc dd Ee
 Aa bb cc DD ee
 aa bb CC DD Ee
 aa bb Cc DD EE
 Aa Bb Cc DD Ee

$$0.3 + 2 + 6 + 0.7 + 0.1 + 0.9$$



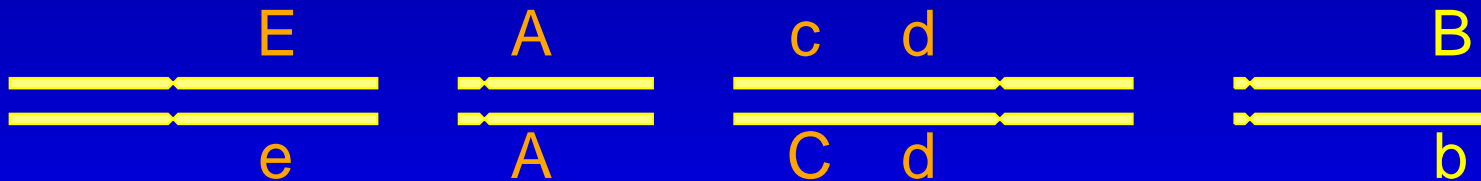
We use the standard quantitative genetic model

$$P = \begin{Bmatrix} AA & 2 \\ Aa & 4 \\ aa & 7 \end{Bmatrix} + \begin{Bmatrix} BB & 0.6 \\ Bb & 0.1 \\ bb & -0.2 \end{Bmatrix} + \begin{Bmatrix} CC & -1 \\ Cc & 6 \\ cc & 6 \end{Bmatrix} + \begin{Bmatrix} DD & 0.3 \\ Dd & 0.3 \\ dd & 0.7 \end{Bmatrix} + \begin{Bmatrix} EE & -0.4 \\ Ee & 0.3 \\ ee & -0.3 \end{Bmatrix} + \text{environmental effect}$$



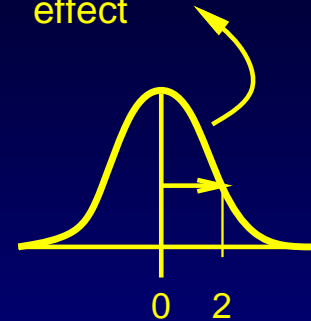
AA Bb Cc dd Ee
 Aa bb cc DD ee
 aa bb CC DD Ee
 aa bb Cc DD EE
 Aa Bb Cc DD Ee

$$0.3 + 2 + 6 + 0.7 + 0.1 + 0.9$$



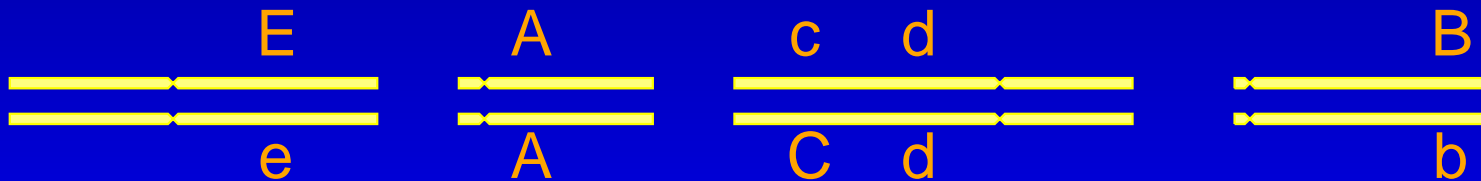
We use the standard quantitative genetic model

$$P = \left\{ \begin{array}{l} AA \ 2 \\ Aa \ 4 \\ aa \ 7 \end{array} \right\} + \left\{ \begin{array}{l} BB \ 0.6 \\ Bb \ 0.1 \\ bb \ -0.2 \end{array} \right\} + \left\{ \begin{array}{l} CC \ -1 \\ Cc \ 6 \\ cc \ 6 \end{array} \right\} + \left\{ \begin{array}{l} DD \ 0.3 \\ Dd \ 0.3 \\ dd \ 0.7 \end{array} \right\} + \left\{ \begin{array}{l} EE \ -0.4 \\ Ee \ 0.3 \\ ee \ -0.3 \end{array} \right\} + \text{environmental effect}$$



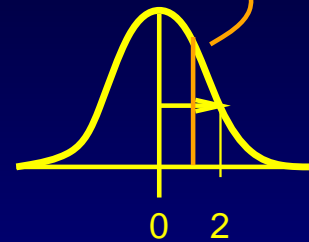
AA Bb Cc dd Ee
 Aa bb cc DD ee
 aa bb CC DD Ee
 aa bb Cc DD EE
 Aa Bb Cc DD Ee

$$0.3 + 2 + 6 + 0.7 + 0.1 + 0.9$$



... which leads to various character values

$$P = \begin{Bmatrix} AA & 2 \\ Aa & 4 \\ aa & 7 \end{Bmatrix} + \begin{Bmatrix} BB & 0.6 \\ Bb & 0.1 \\ bb & -0.2 \end{Bmatrix} + \begin{Bmatrix} CC & -1 \\ Cc & 6 \\ cc & 6 \end{Bmatrix} + \begin{Bmatrix} DD & 0.3 \\ Dd & 0.3 \\ dd & 0.7 \end{Bmatrix} + \begin{Bmatrix} EE & -0.4 \\ Ee & 0.3 \\ ee & -0.3 \end{Bmatrix} + \text{environmental effect}$$



AA Bb Cc dd Ee 10

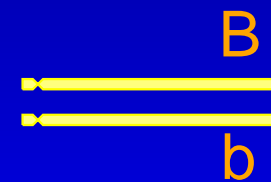
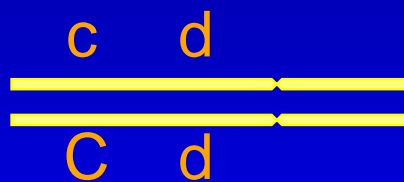
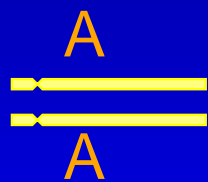
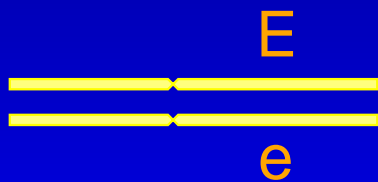
Aa bb cc DD ee

aa bb CC DD Ee

aa bb Cc DD EE

Aa Bb Cc DD Ee

$$0.3 + 2 + 6 + 0.7 + 0.1 + 0.9$$



... which leads to various character values

$$P = \begin{Bmatrix} AA & 2 \\ Aa & 4 \\ aa & 7 \end{Bmatrix} + \begin{Bmatrix} BB & 0.6 \\ Bb & 0.1 \\ bb & -0.2 \end{Bmatrix} + \begin{Bmatrix} CC & -1 \\ Cc & 6 \\ cc & 6 \end{Bmatrix} + \begin{Bmatrix} DD & 0.3 \\ Dd & 0.3 \\ dd & 0.7 \end{Bmatrix} + \begin{Bmatrix} EE & -0.4 \\ Ee & 0.3 \\ ee & -0.3 \end{Bmatrix} + \text{environmental effect}$$

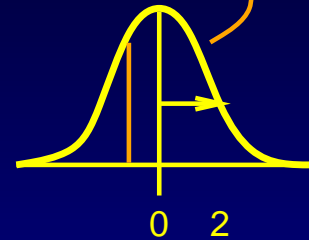
AA Bb Cc dd Ee 10

Aa bb cc DD ee 8.8

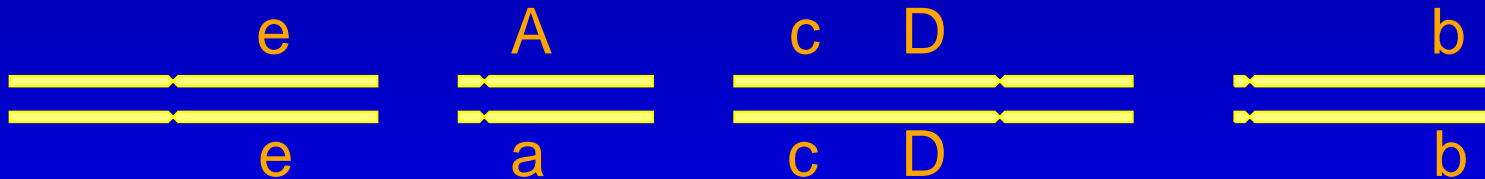
aa bb CC DD Ee

aa bb Cc DD EE

Aa Bb Cc DD Ee

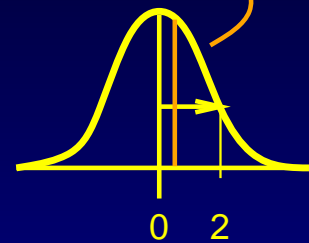


$$-0.3 + 4 + 6 + 0.3 \quad -0.2 - 1.0$$



... which leads to various character values

$$P = \begin{Bmatrix} AA & 2 \\ Aa & 4 \\ aa & 7 \end{Bmatrix} + \begin{Bmatrix} BB & 0.6 \\ Bb & 0.1 \\ bb & -0.2 \end{Bmatrix} + \begin{Bmatrix} CC & -1 \\ Cc & 6 \\ cc & 6 \end{Bmatrix} + \begin{Bmatrix} DD & 0.3 \\ Dd & 0.3 \\ dd & 0.7 \end{Bmatrix} + \begin{Bmatrix} EE & -0.4 \\ Ee & 0.3 \\ ee & -0.3 \end{Bmatrix} + \text{environmental effect}$$



AA Bb Cc dd Ee 10

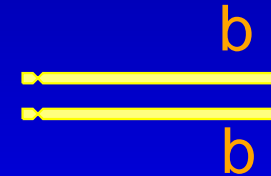
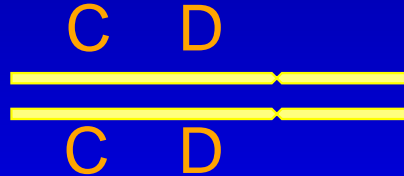
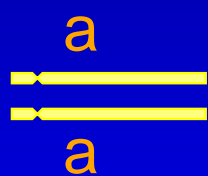
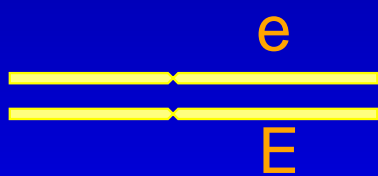
Aa bb cc DD ee 8.8

aa bb CC DD Ee 6.5

aa bb Cc DD EE

Aa Bb Cc DD Ee

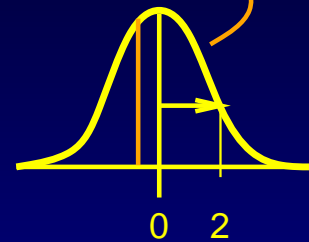
$$0.3 + 7 + -1 + 0.3 + -0.2 + 0.1$$



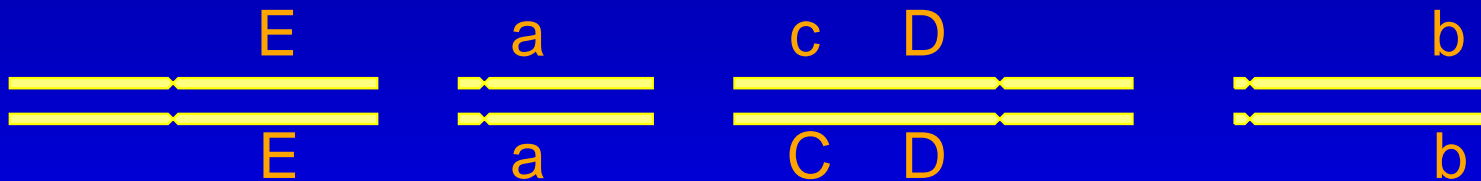
... which leads to various character values

$$P = \begin{Bmatrix} AA & 2 \\ Aa & 4 \\ aa & 7 \end{Bmatrix} + \begin{Bmatrix} BB & 0.6 \\ Bb & 0.1 \\ bb & -0.2 \end{Bmatrix} + \begin{Bmatrix} CC & -1 \\ Cc & 6 \\ cc & 6 \end{Bmatrix} + \begin{Bmatrix} DD & 0.3 \\ Dd & 0.3 \\ dd & 0.7 \end{Bmatrix} + \begin{Bmatrix} EE & -0.4 \\ Ee & 0.3 \\ ee & -0.3 \end{Bmatrix} + \text{environmental effect}$$

AA	Bb	Cc	dd	Ee	10
Aa	bb	cc	DD	ee	8.8
aa	bb	CC	DD	Ee	6.5
aa	bb	Cc	DD	EE	12.2
Aa	Bb	Cc	DD	Ee	

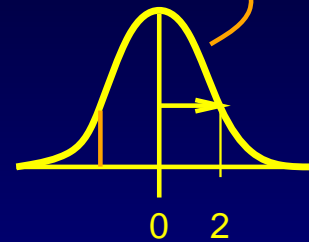


$$-0.4 + 7 + 6 + 0.3 - 0.2 - 0.5$$



... which leads to various character values

$$P = \begin{Bmatrix} AA & 2 \\ Aa & 4 \\ aa & 7 \end{Bmatrix} + \begin{Bmatrix} BB & 0.6 \\ Bb & 0.1 \\ bb & -0.2 \end{Bmatrix} + \begin{Bmatrix} CC & -1 \\ Cc & 6 \\ cc & 6 \end{Bmatrix} + \begin{Bmatrix} DD & 0.3 \\ Dd & 0.3 \\ dd & 0.7 \end{Bmatrix} + \begin{Bmatrix} EE & -0.4 \\ Ee & 0.3 \\ ee & -0.3 \end{Bmatrix} + \text{environmental effect}$$



$$AA \quad Bb \quad Cc \quad dd \quad Ee \quad 10$$

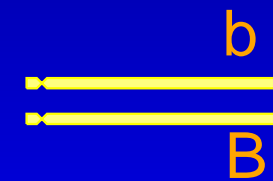
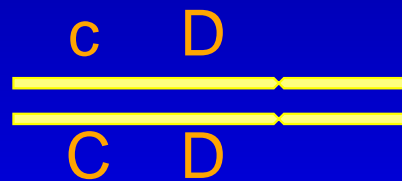
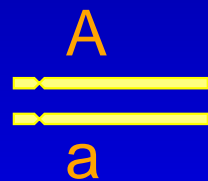
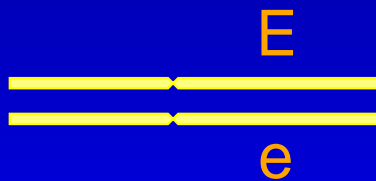
$$Aa \quad bb \quad cc \quad DD \quad ee \quad 8.8$$

$$aa \quad bb \quad CC \quad DD \quad Ee \quad 6.5$$

$$aa \quad bb \quad Cc \quad DD \quad EE \quad 12.2$$

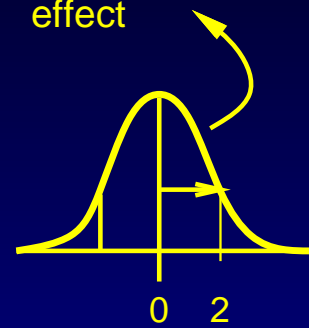
$$Aa \quad Bb \quad Cc \quad DD \quad Ee \quad 8.9$$

$$0.3 + 4 + 6 + 0.3 + 0.1 - 1.8$$



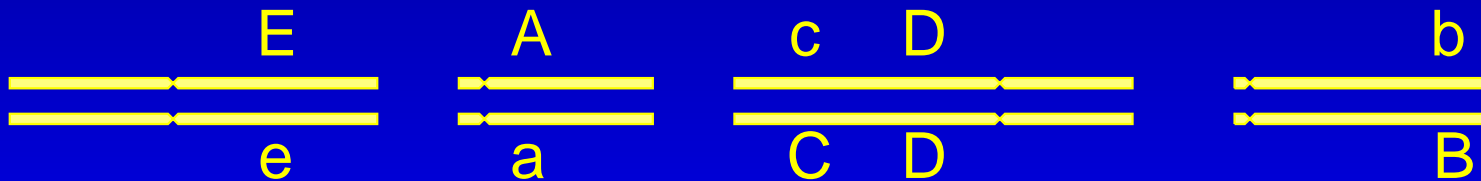
... then we apply a threshold of 9

$$P = \begin{Bmatrix} AA & 2 \\ Aa & 4 \\ aa & 7 \end{Bmatrix} + \begin{Bmatrix} BB & 0.6 \\ Bb & 0.1 \\ bb & -0.2 \end{Bmatrix} + \begin{Bmatrix} CC & -1 \\ Cc & 6 \\ cc & 6 \end{Bmatrix} + \begin{Bmatrix} DD & 0.3 \\ Dd & 0.3 \\ dd & 0.7 \end{Bmatrix} + \begin{Bmatrix} EE & -0.4 \\ Ee & 0.3 \\ ee & -0.3 \end{Bmatrix} + \text{environmental effect}$$



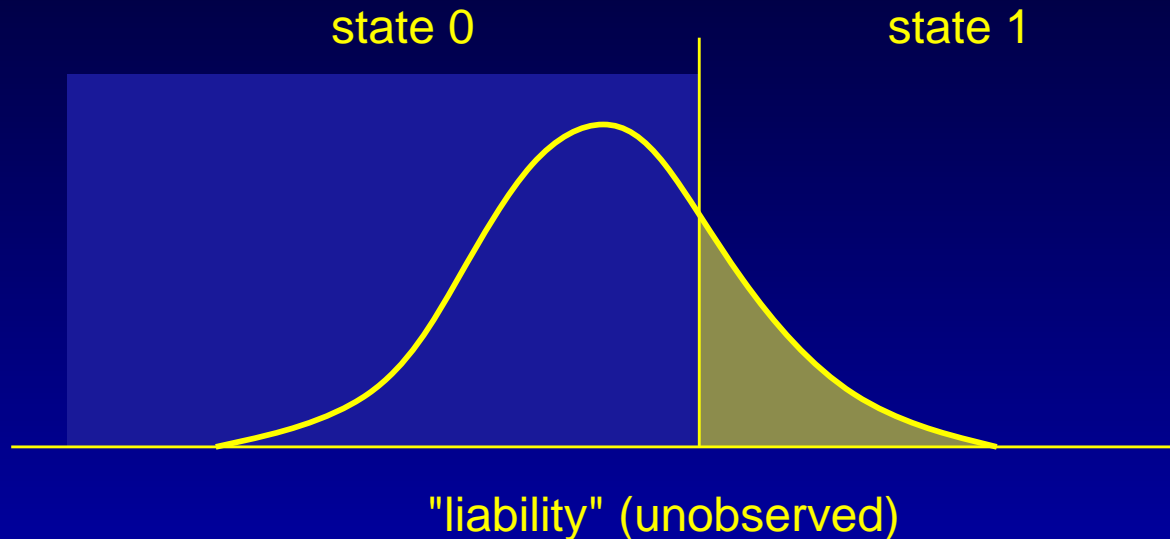
AA	Bb	Cc	dd	Ee	10	→	1
Aa	bb	cc	DD	ee	8.8	→	0
aa	bb	CC	DD	Ee	6.5	→	0
aa	bb	Cc	DD	EE	12.2	→	1
Aa	Bb	Cc	DD	Ee	8.9	→	0

$$0.3 + 4 + 6 + 0.3 + 0.1 - 1.8$$



The threshold model

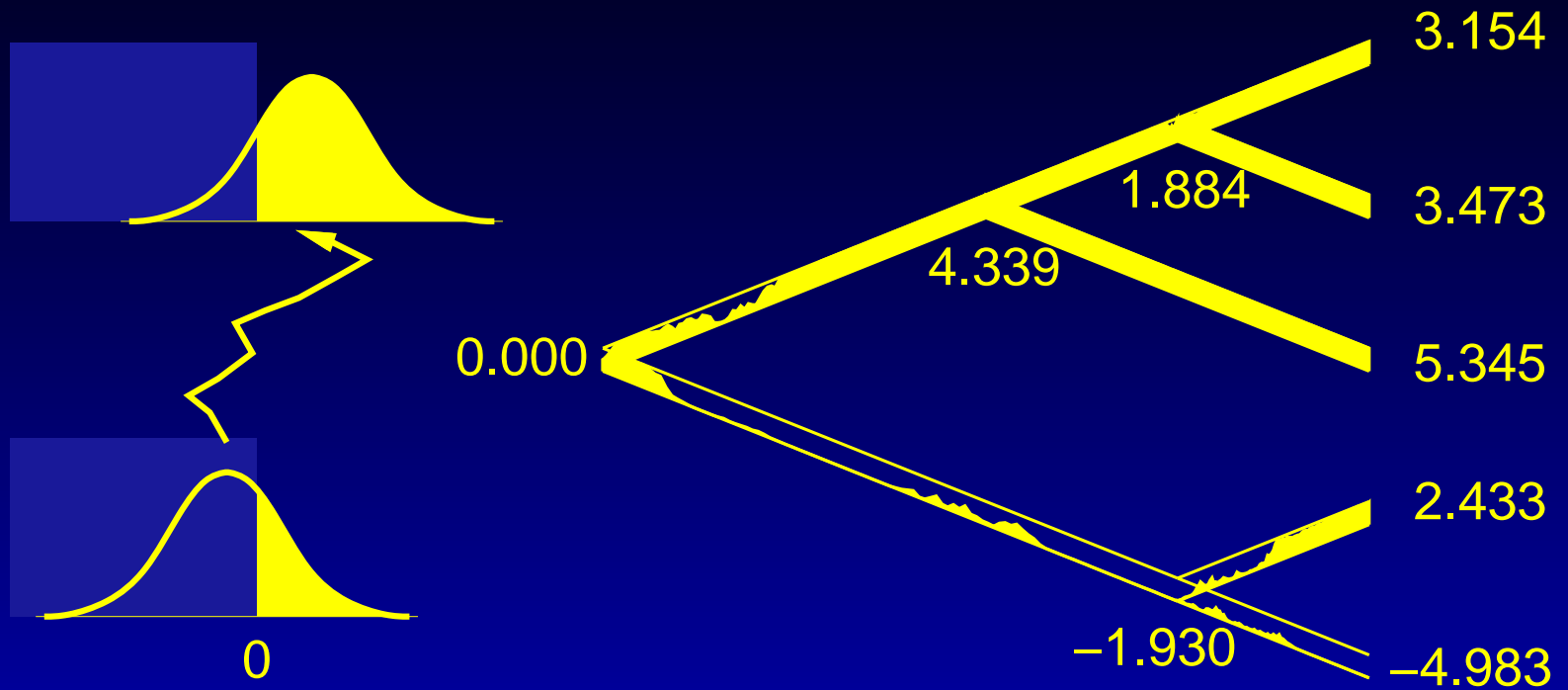
The threshold model (Wright, 1934; Falconer, 1965), plus Brownian motion



Advantages:

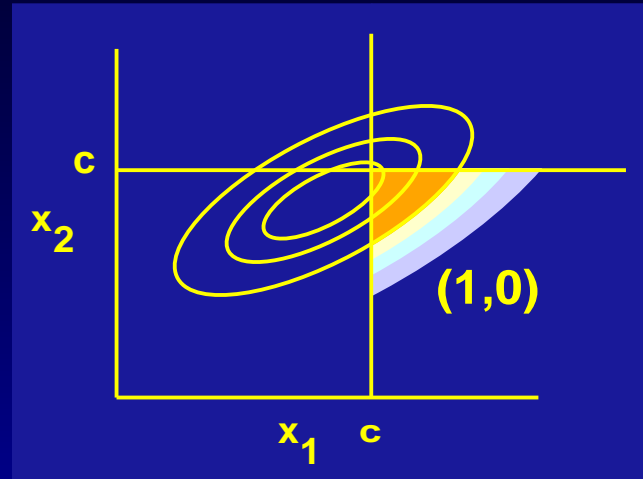
1. Predicts polymorphism as a lineage crosses the threshold
2. Soon after the threshold is crossed, one is more likely to revert. Less later.
3. Can allow covariation of characters

The threshold model on a tree



Computing the likelihood

With two species, one character:



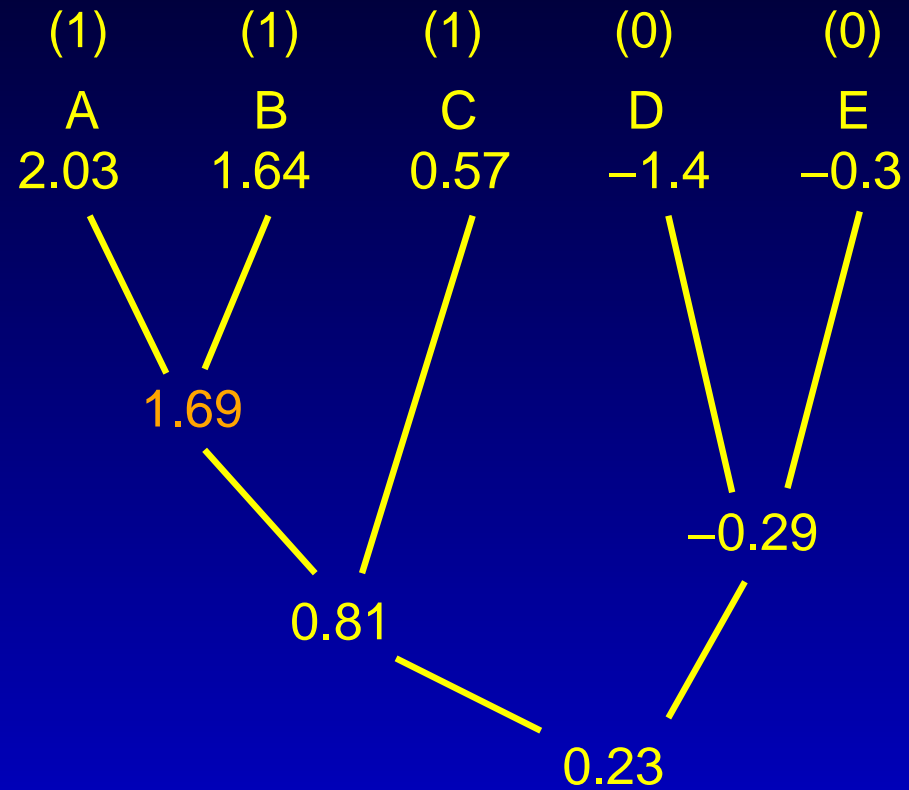
Disadvantages:

Quite hard to compute likelihoods: need to compute area in a corner of a correlated multivariate normal distribution.

With 5 species, one character:

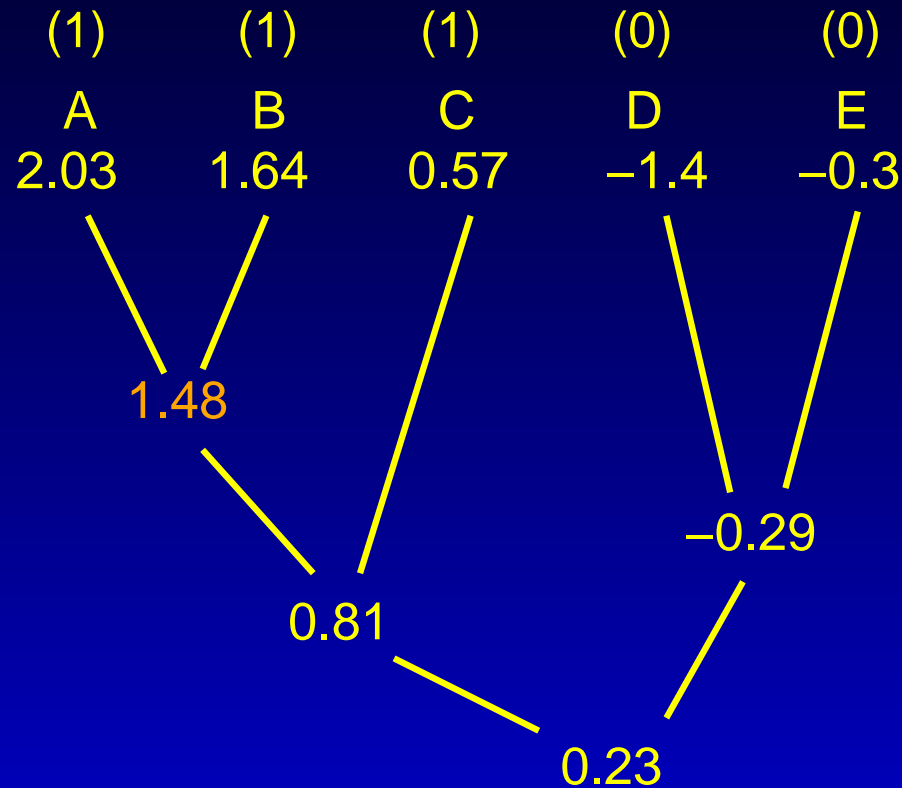
$$\begin{aligned} L &= \text{Prob}(1, 1, 0, 1, 1) \\ &= \int_0^\infty \int_0^\infty \int_{-\infty}^0 \int_0^\infty \int_0^\infty \varphi(x_1, x_2, x_3, x_4, x_5 \mid \text{Tree}) dx_1 dx_2 dx_3 dx_4 dx_5 \end{aligned}$$

MCMC on liabilities

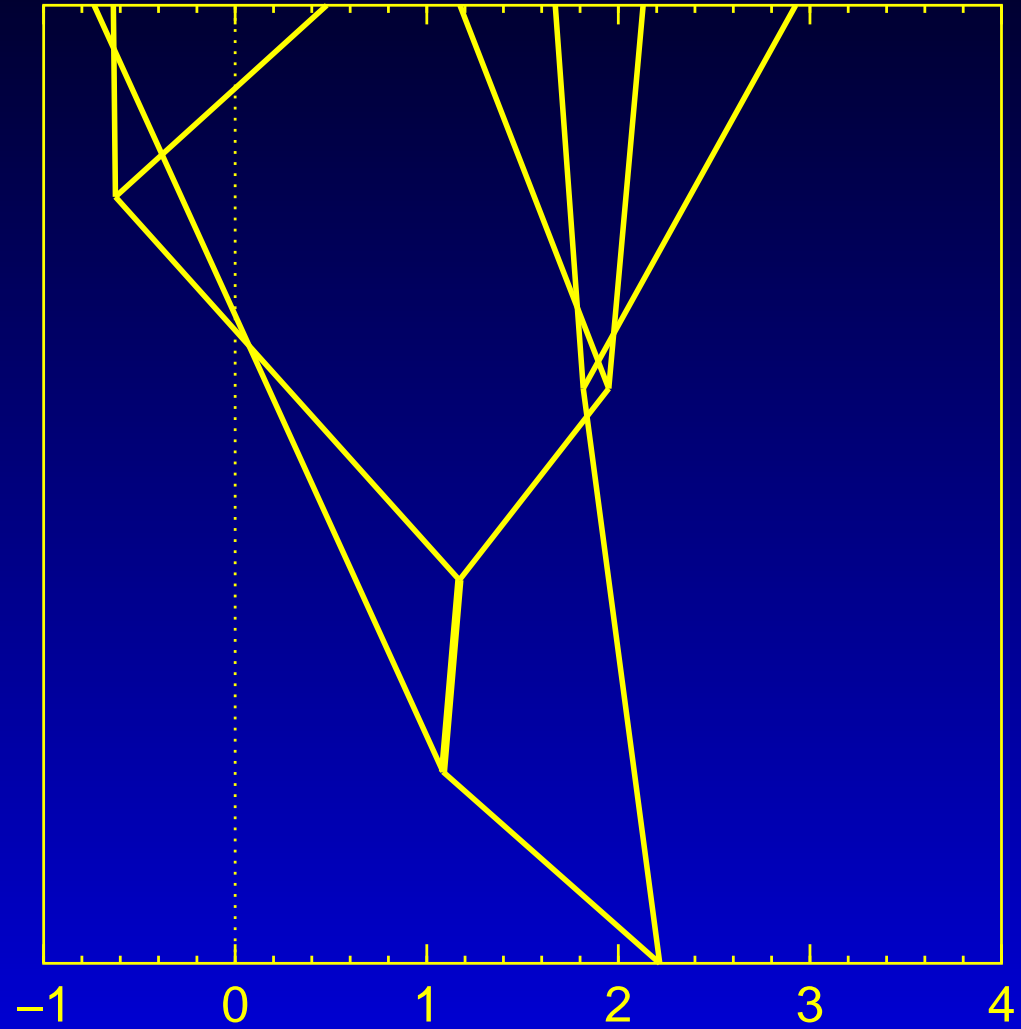


MCMC on liabilities: result of Gibbs sampling

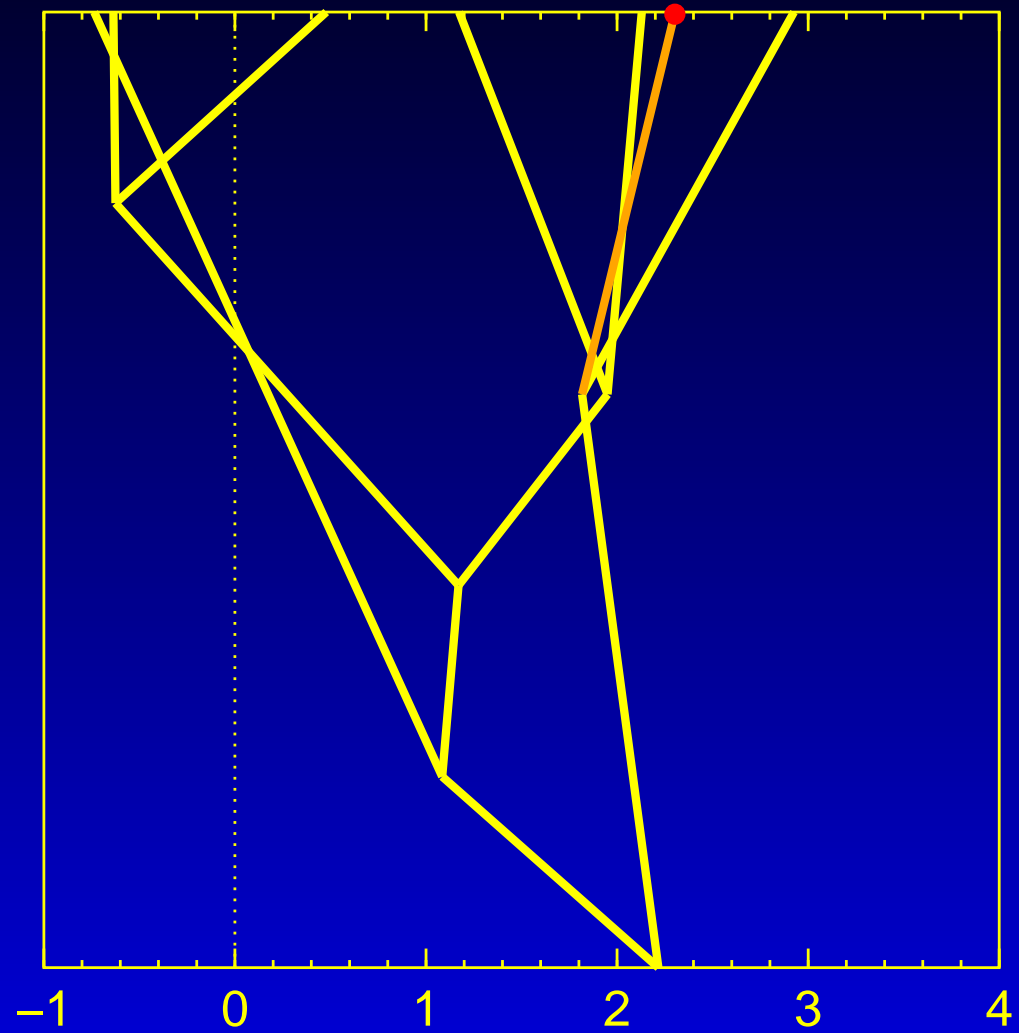
Gibbs sampler for internal node values



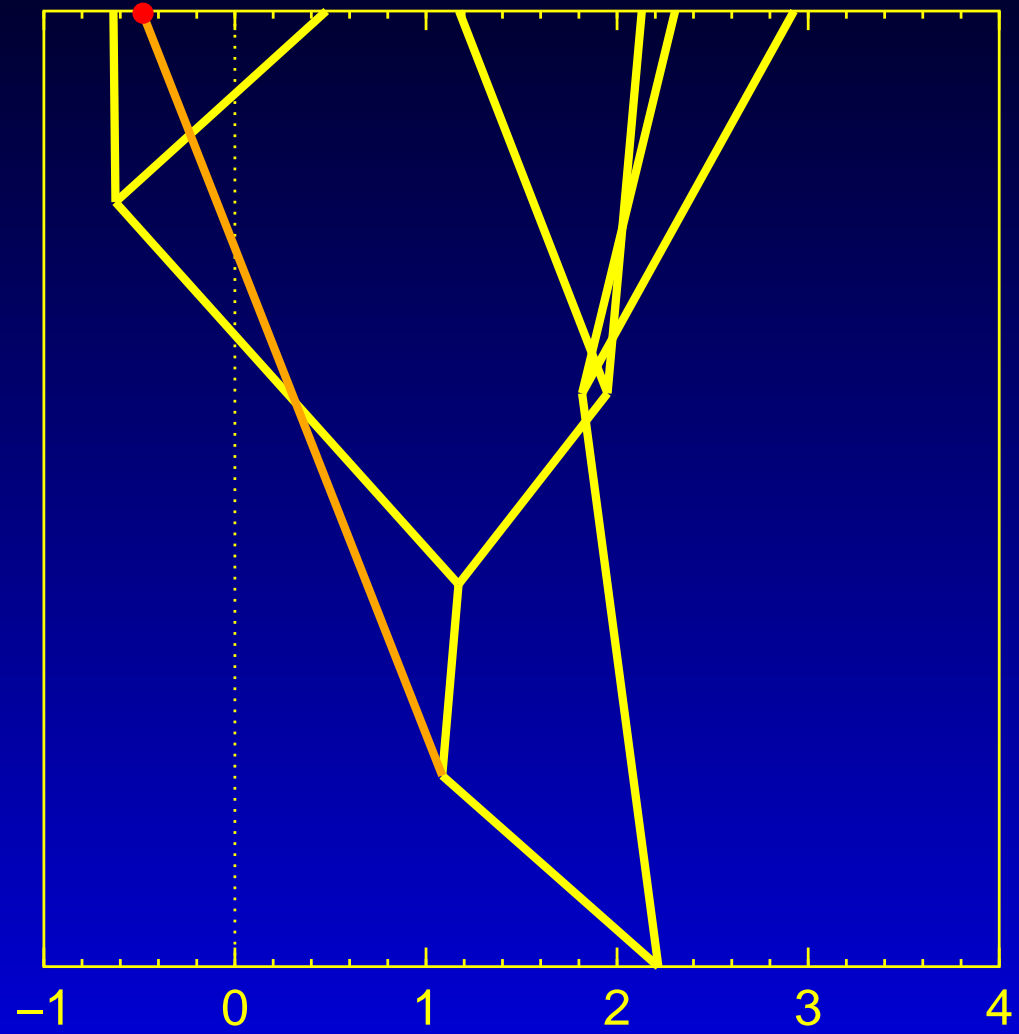
An example



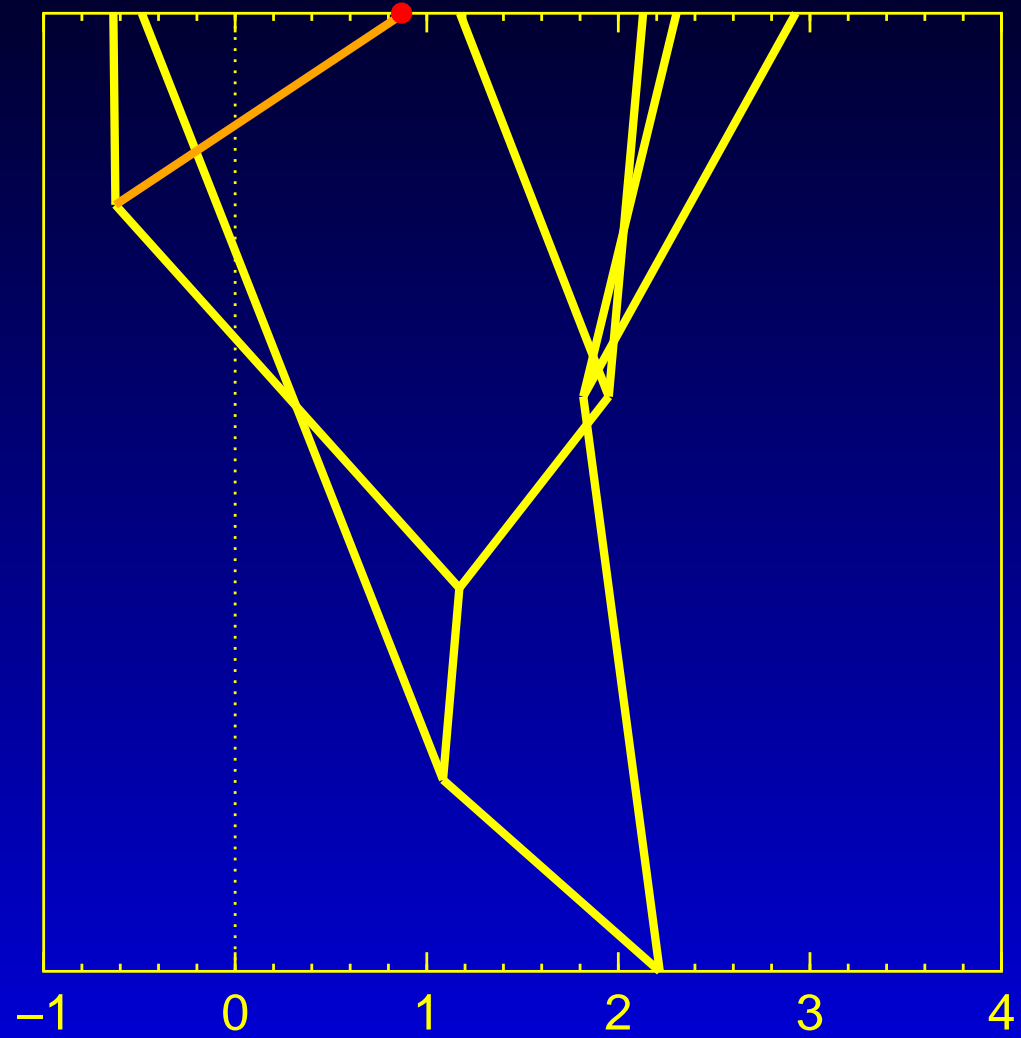
An example



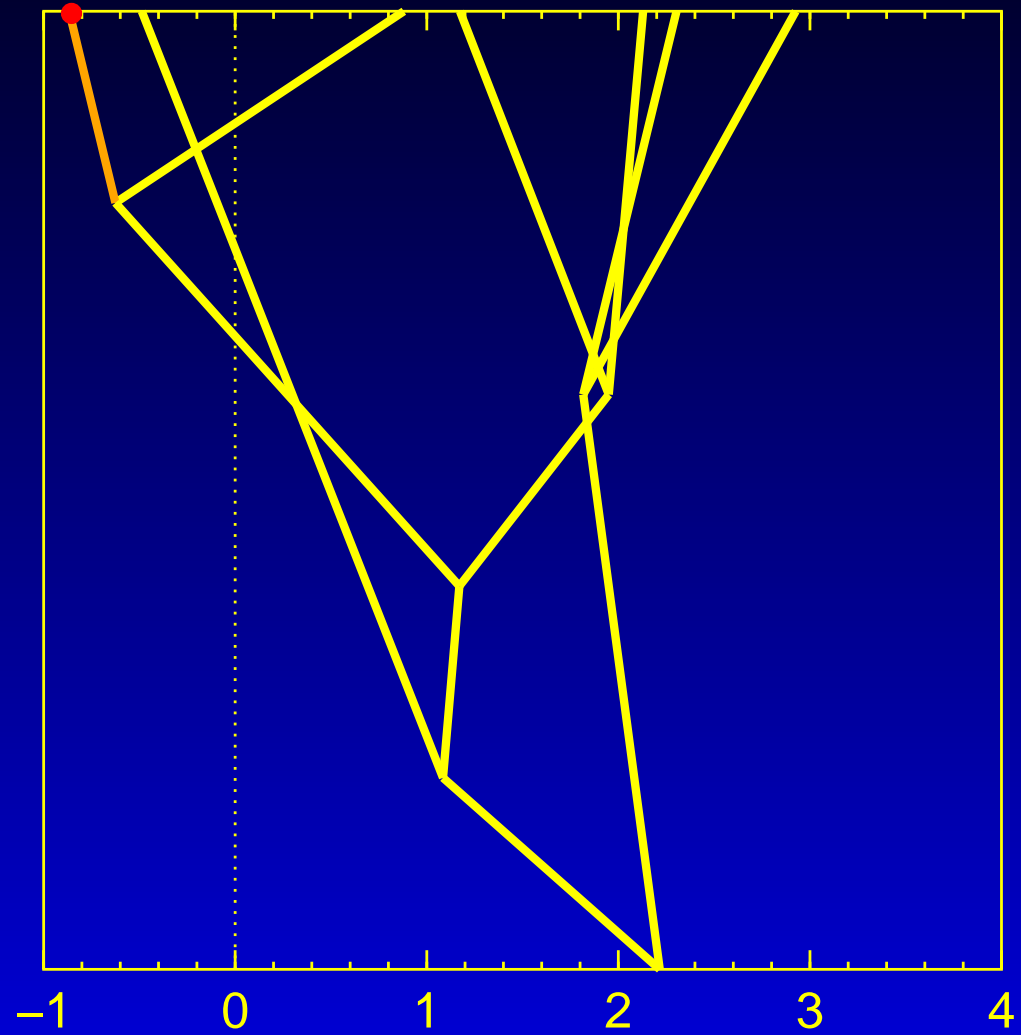
An example



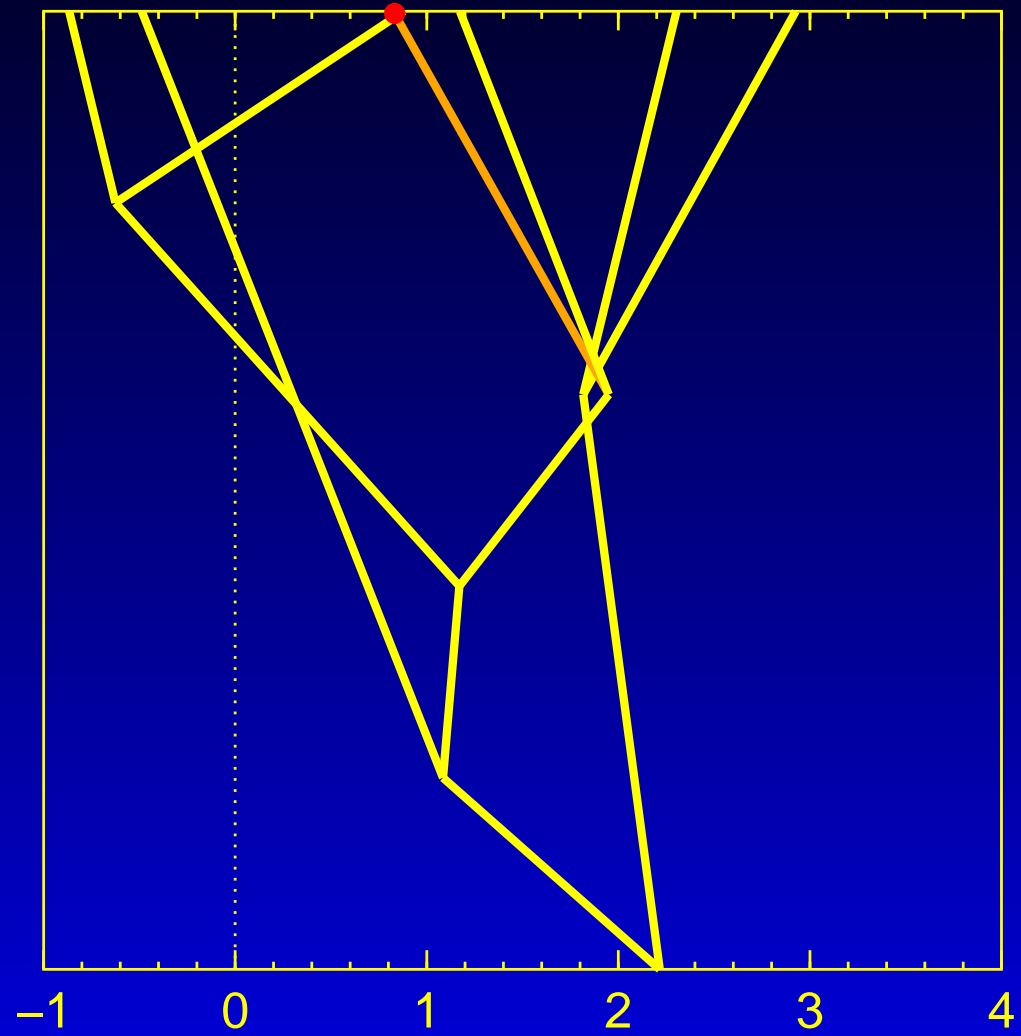
An example



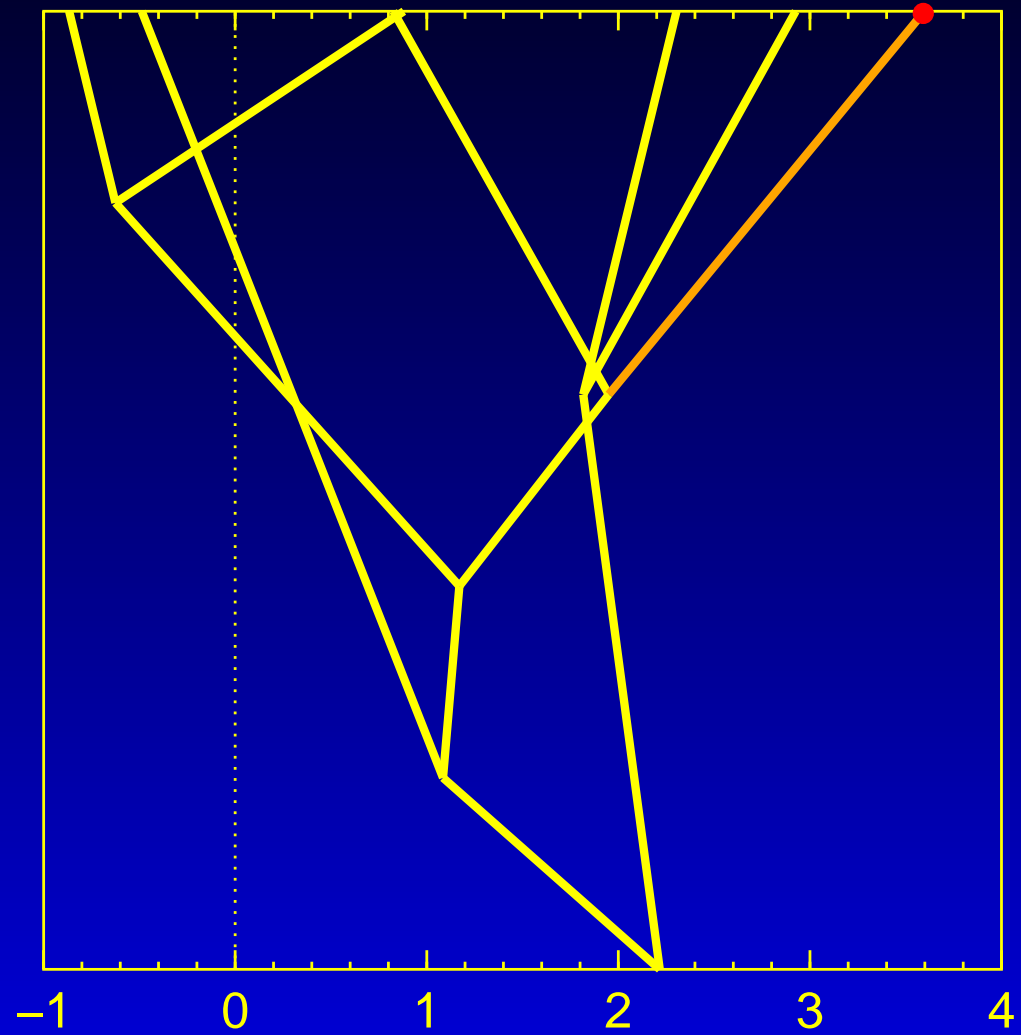
An example



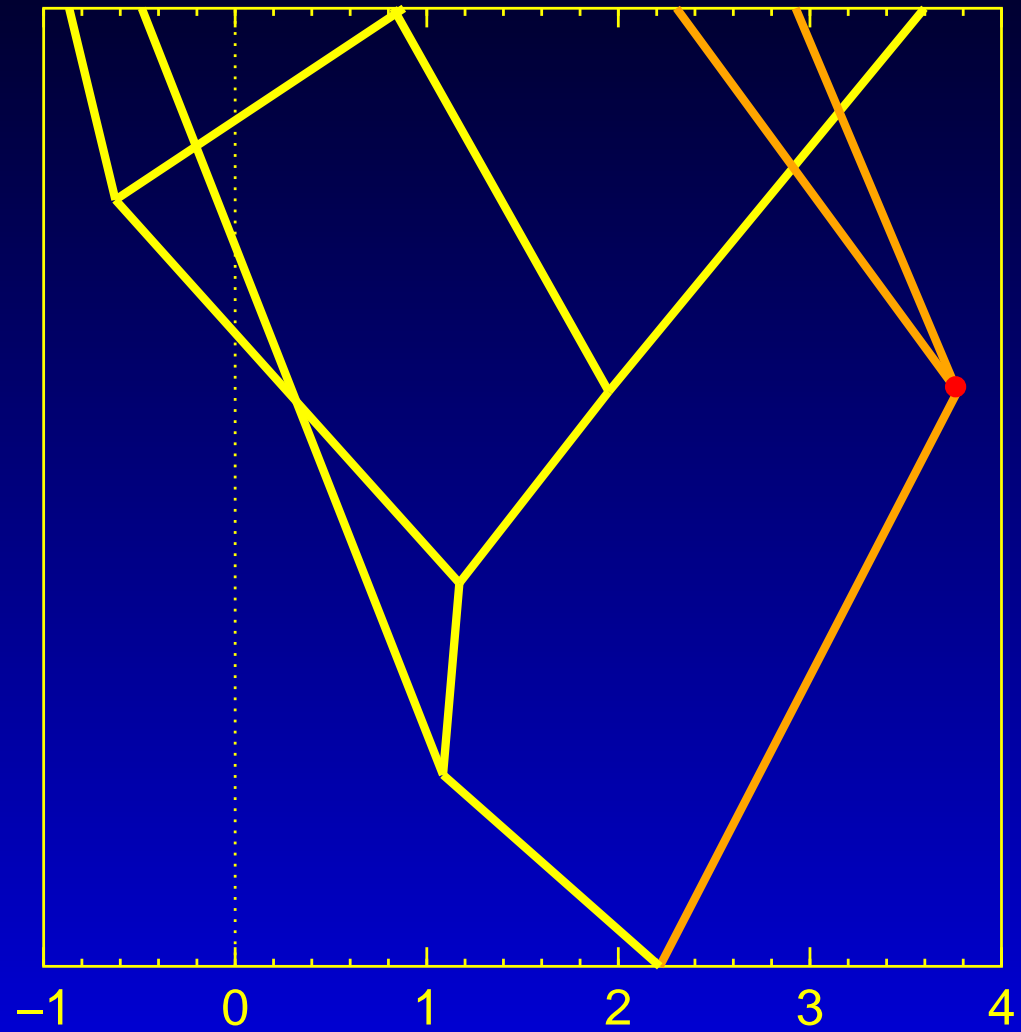
An example



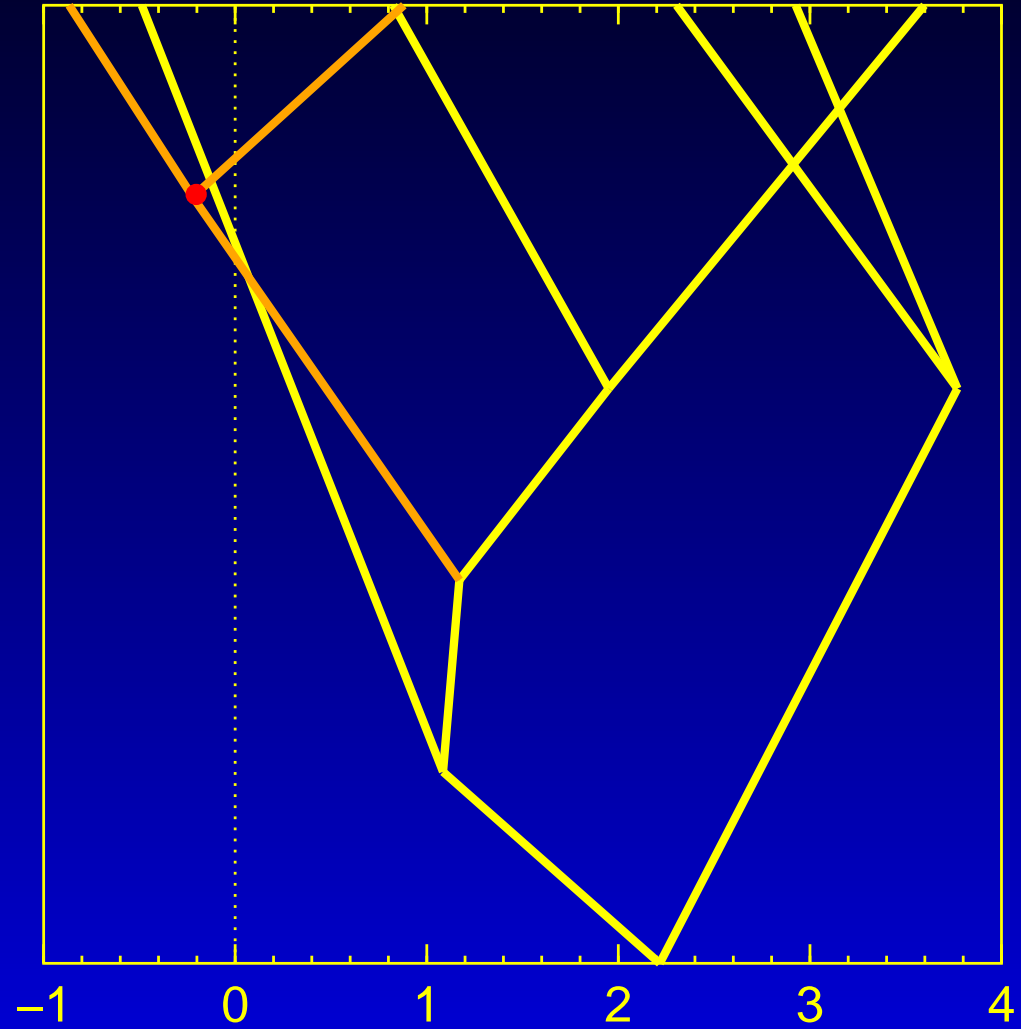
An example



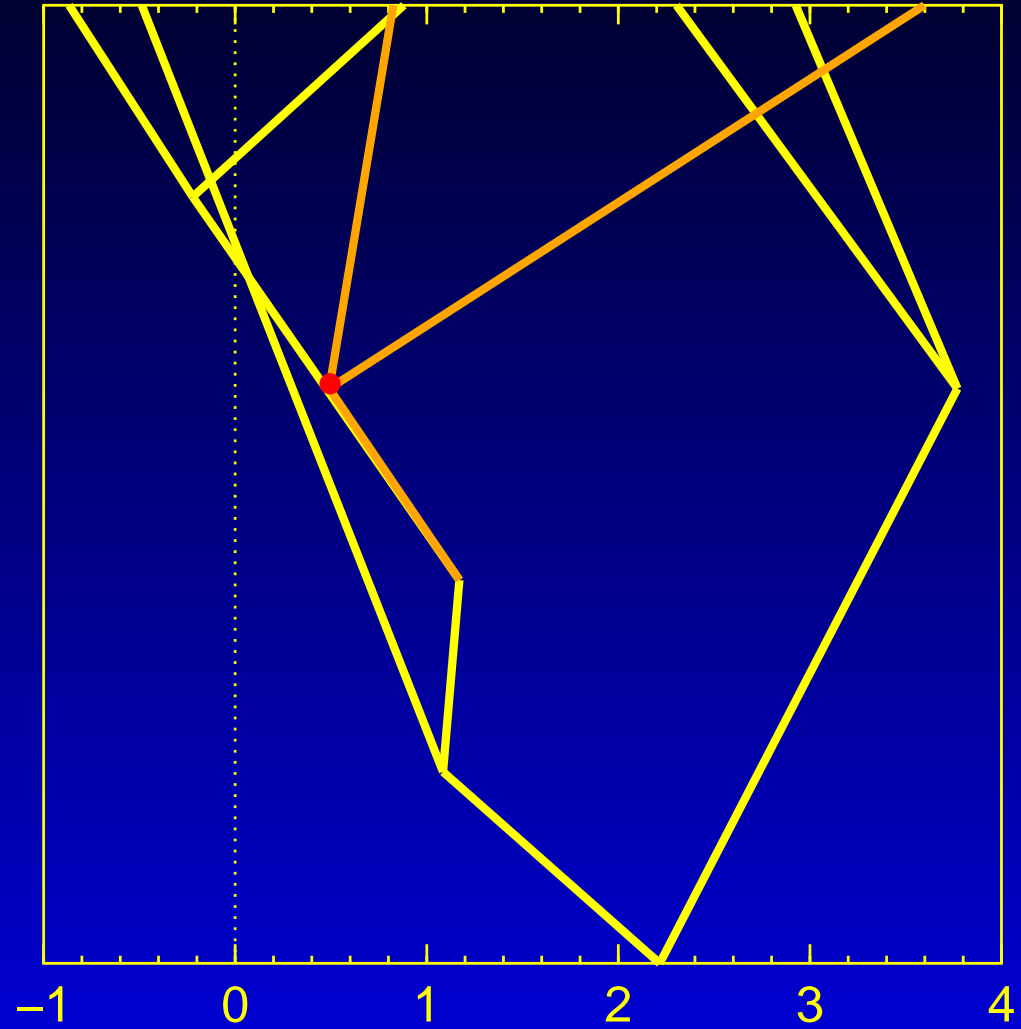
An example



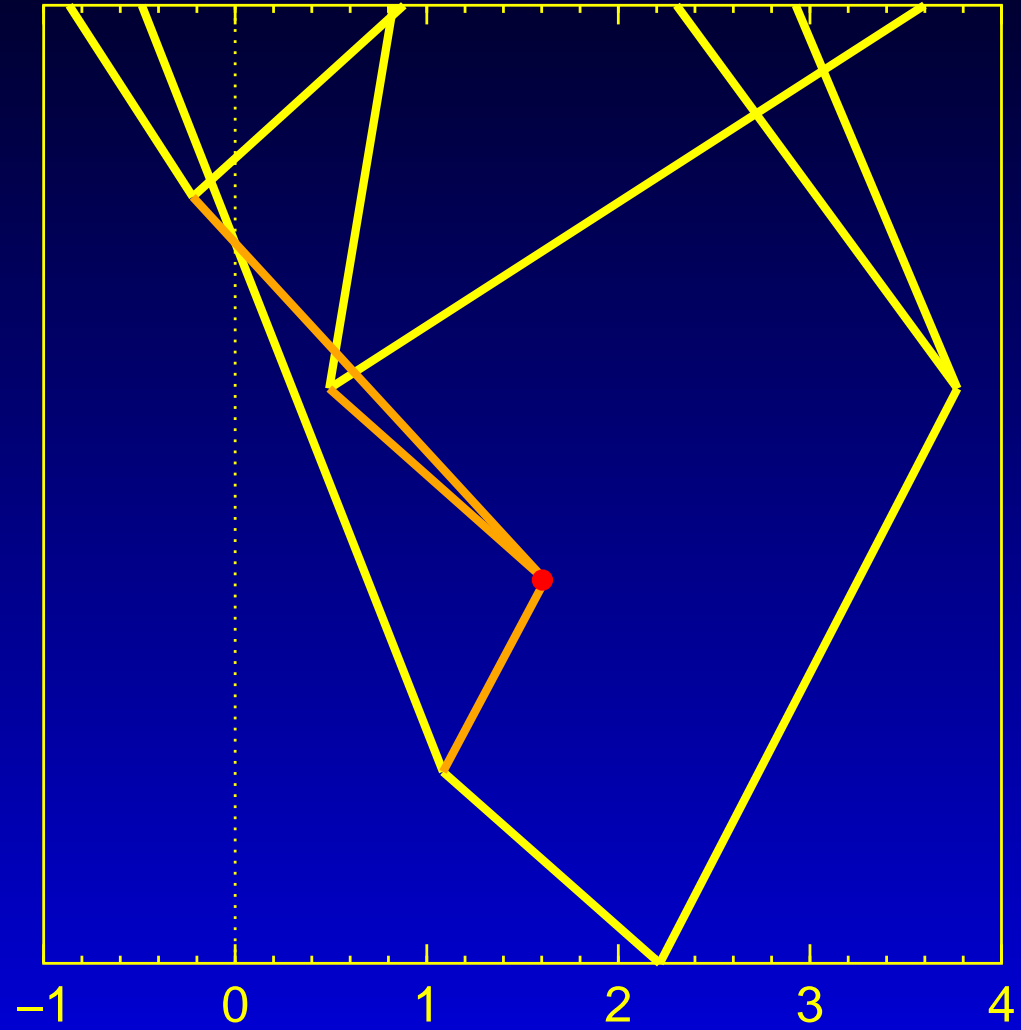
An example



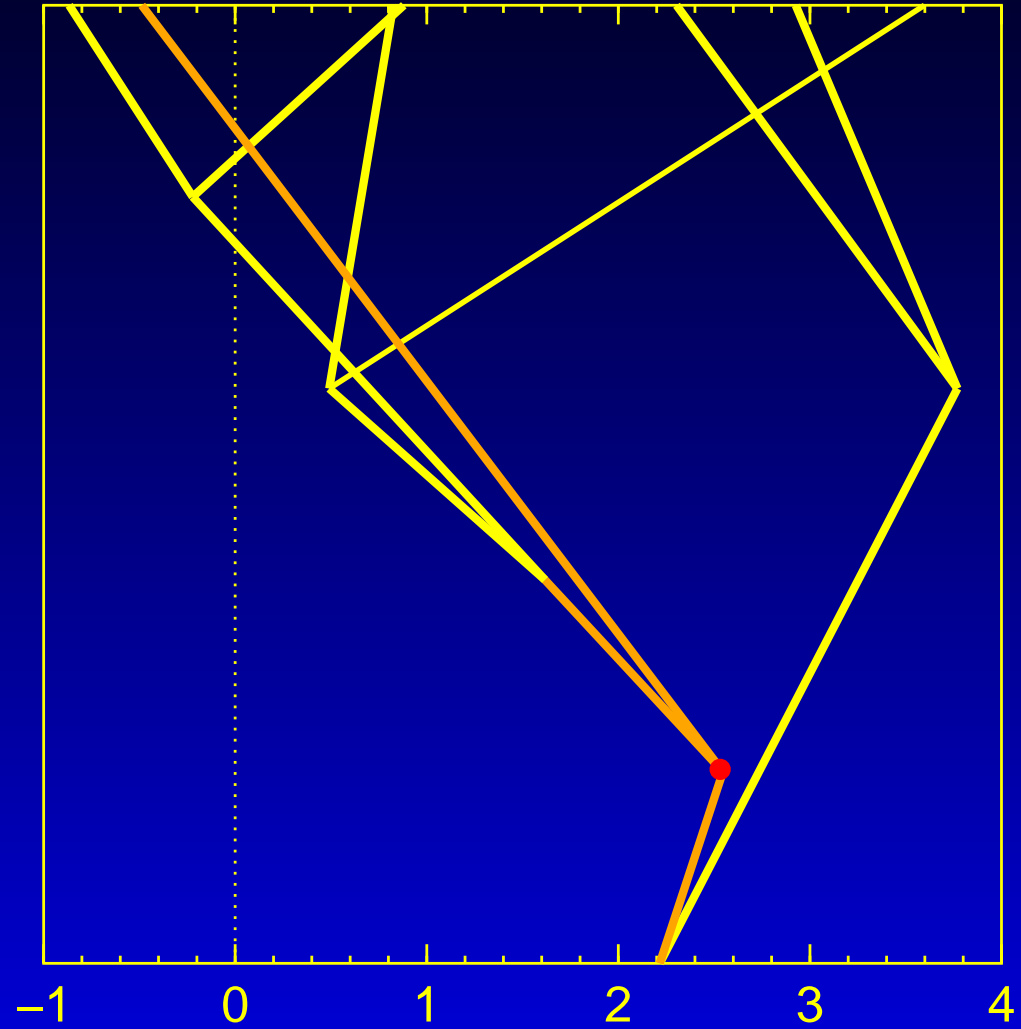
An example



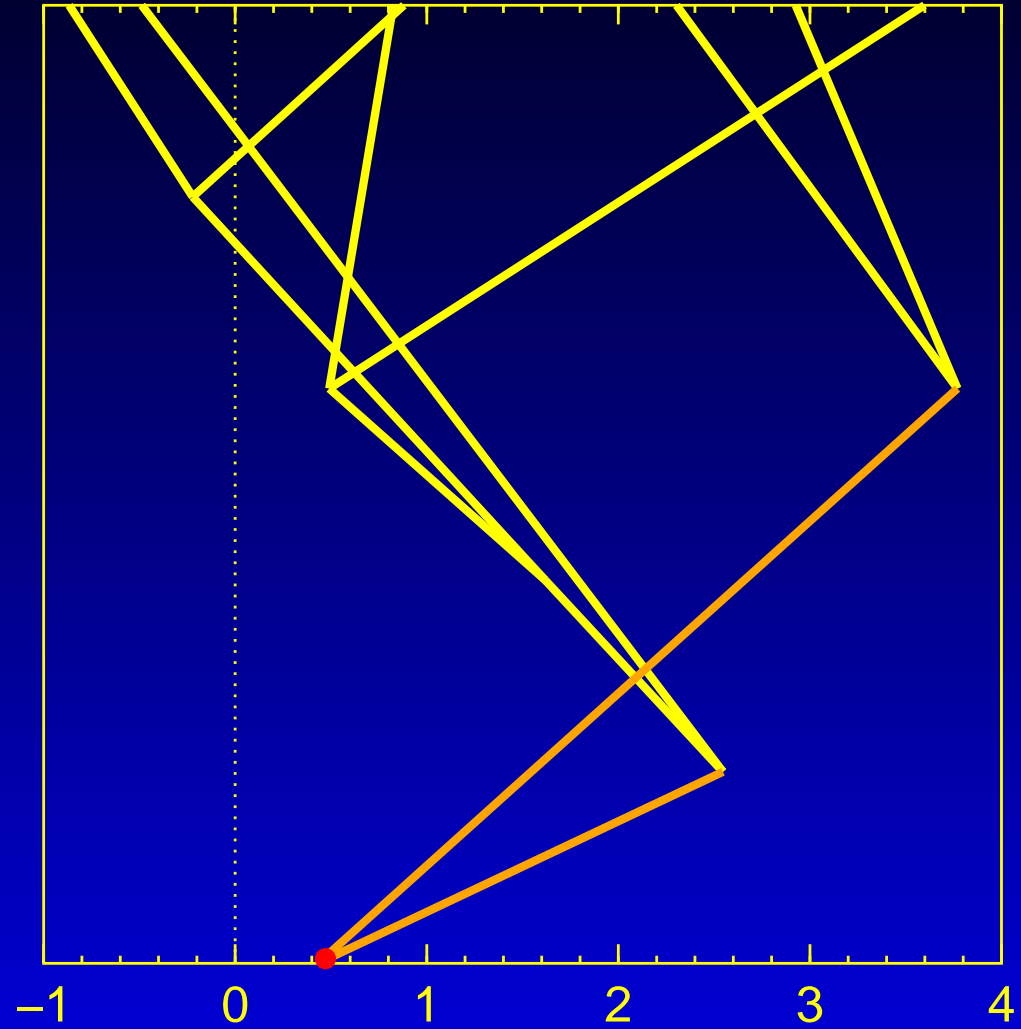
An example



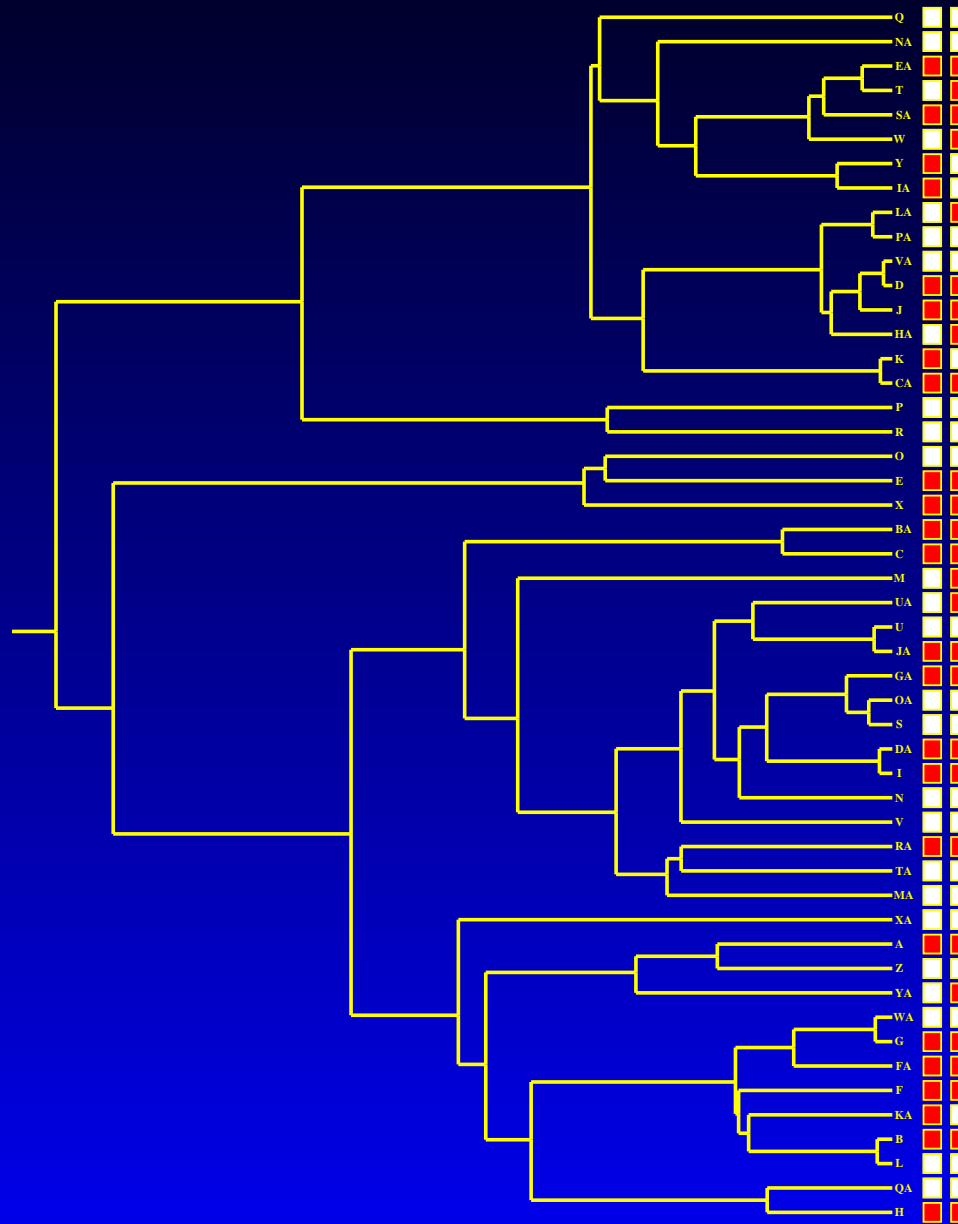
An example



An example



A simulated example, with its true tree



A (very) small-scale simulation

True covariance matrix

1	0.8	0
0.8	1.64	-0.6
0	-0.6	1.36

With 50 species on the given tree:

For the same simulated data set

	Run 1			Run 2		
1.70072	1.39356	0.39289	1.68167	1.40056	0.40495	
1.39356	1.65104	0.20923	1.40056	1.67836	0.23021	
0.39289	0.20923	1.08066	0.40495	0.23021	1.09550	

Continuous characters too? Both?

- We could do this for continuous characters too by assuming the tip liabilities are observed and don't change

Continuous characters too? Both?

- We could do this for continuous characters too by assuming the tip liabilities are observed and don't change
- Then we could estimate covariances.

Continuous characters too? Both?

- We could do this for continuous characters too by assuming the tip liabilities are observed and don't change
- Then we could estimate covariances.
- This would be a noisier version of what we can do exactly already, so there is no point to it, right?

Continuous characters too? Both?

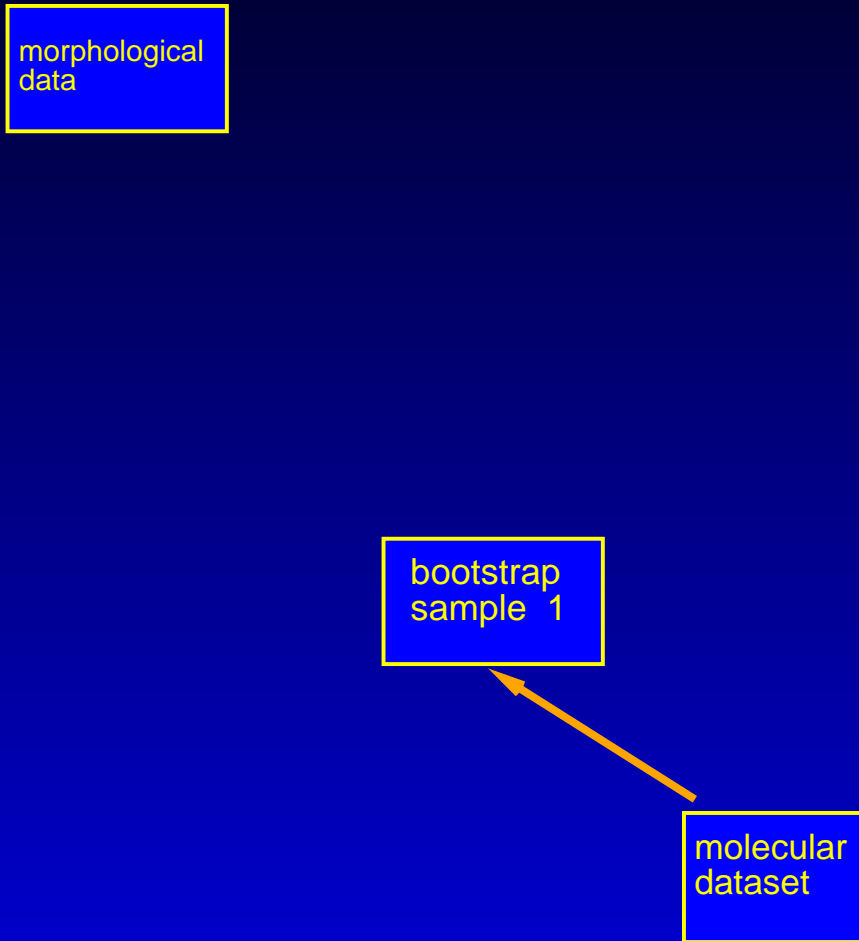
- We could do this for continuous characters too by assuming the tip liabilities are observed and don't change
- Then we could estimate covariances.
- This would be a noisier version of what we can do exactly already, so there is no point to it, right?
- But ... we could do both discrete and continuous characters together in this way, with almost no extra effort.

When the tree is noisy: Propagating bootstrap sampling

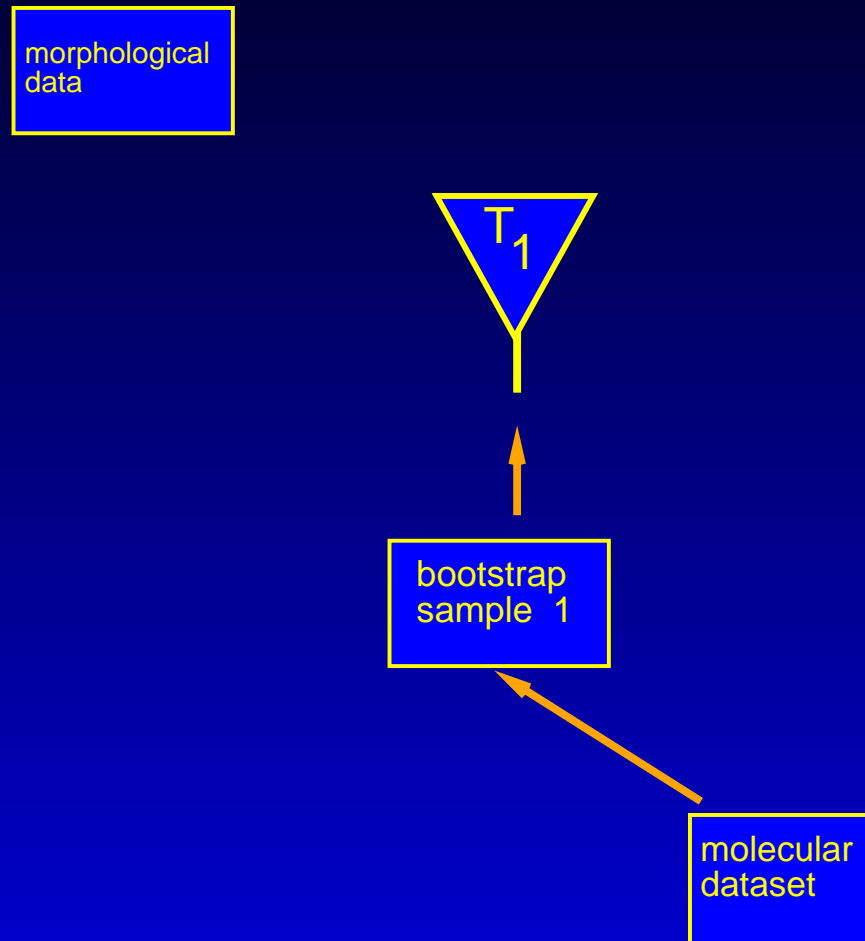
morphological
data

molecular
dataset

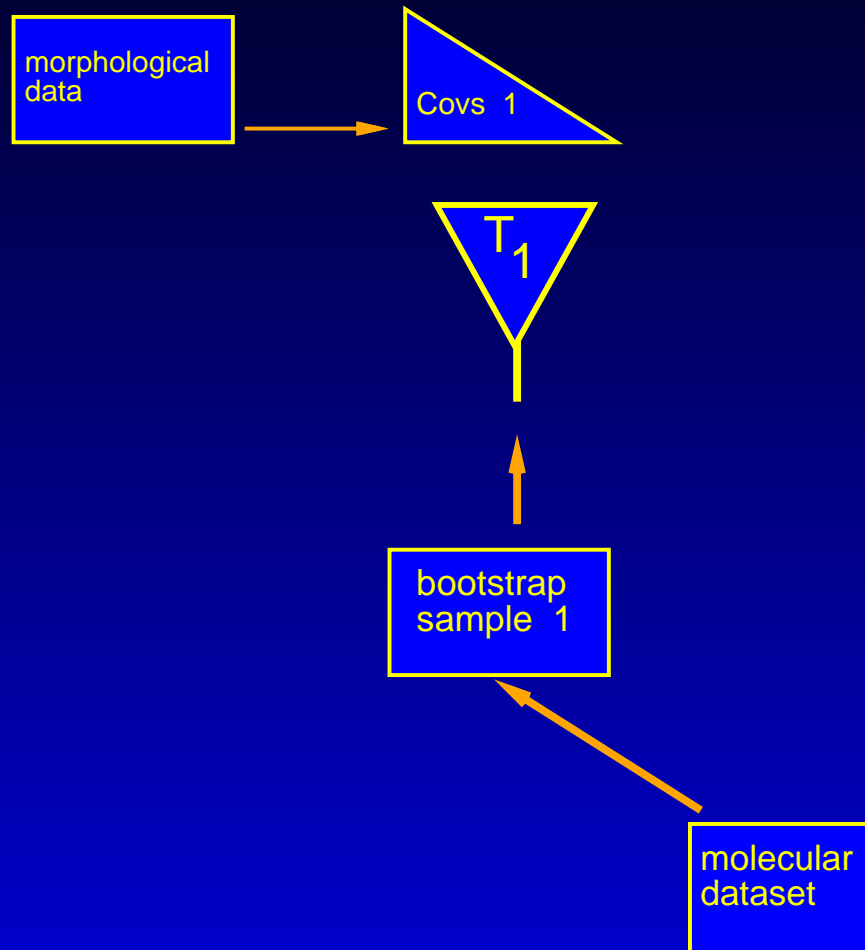
Propagating bootstrap sampling



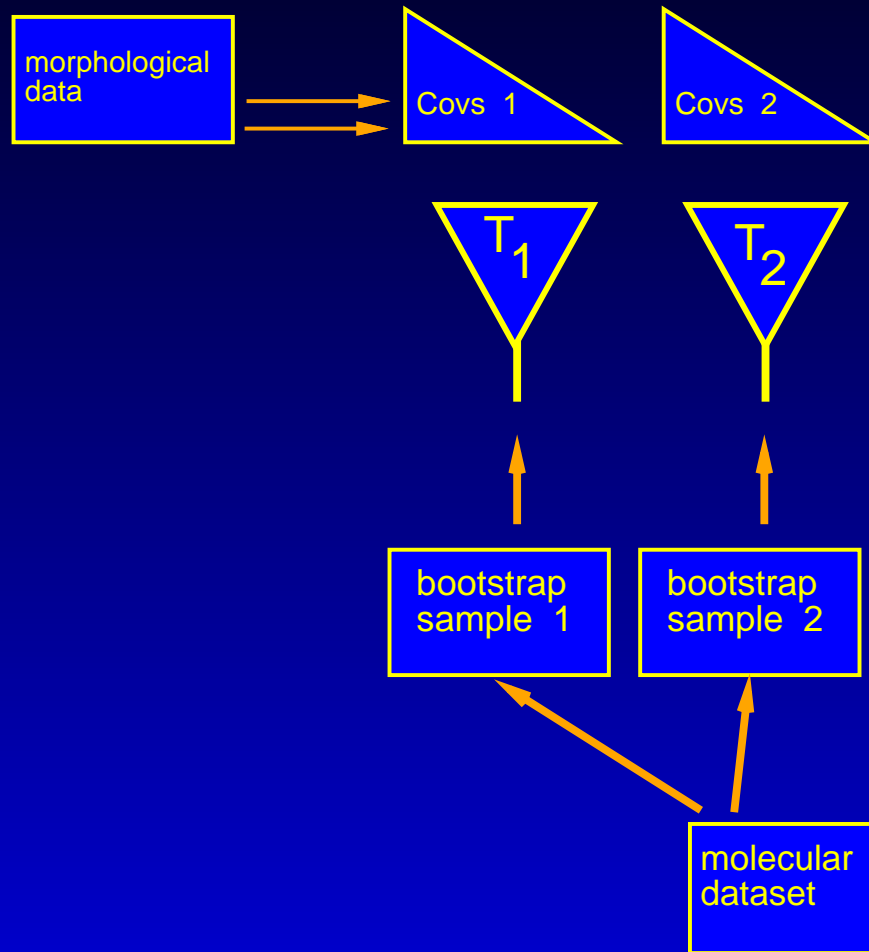
Propagating bootstrap sampling



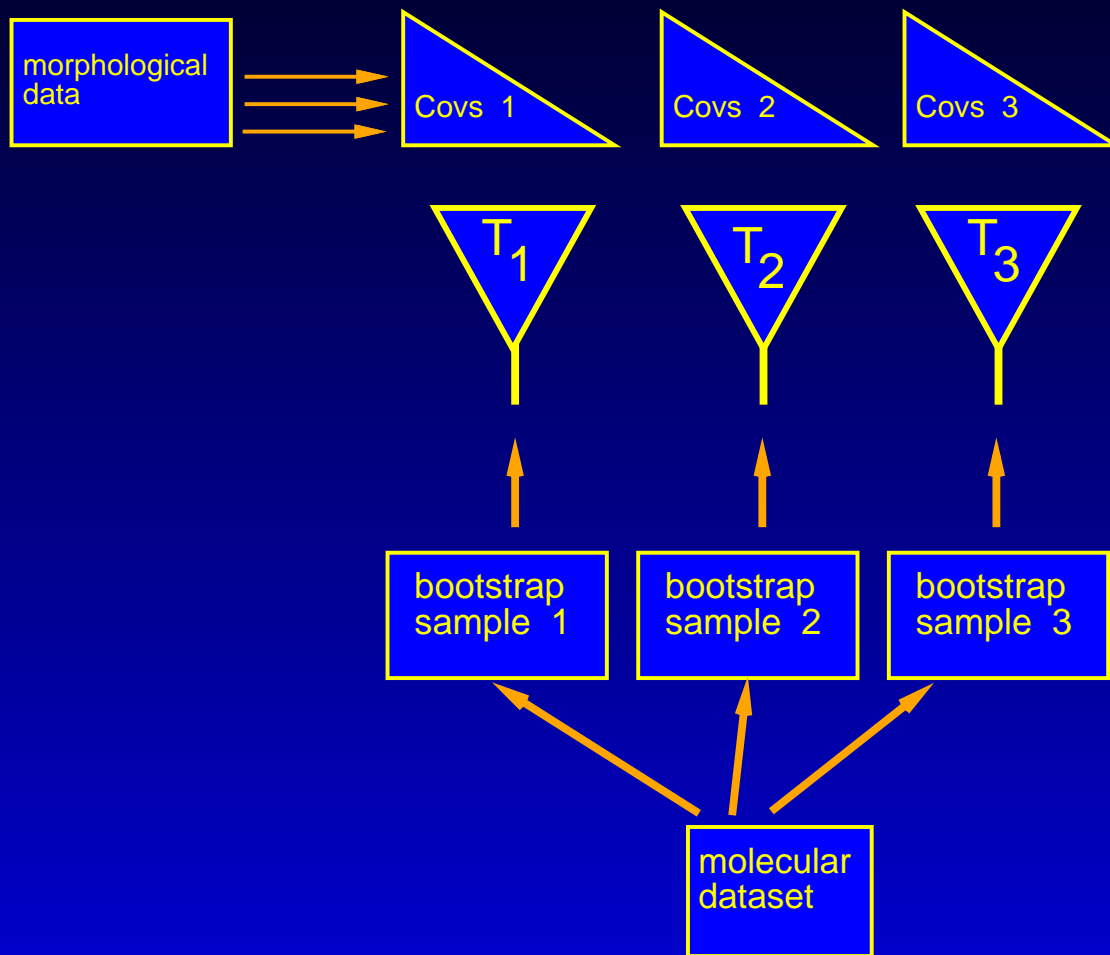
Propagating bootstrap sampling



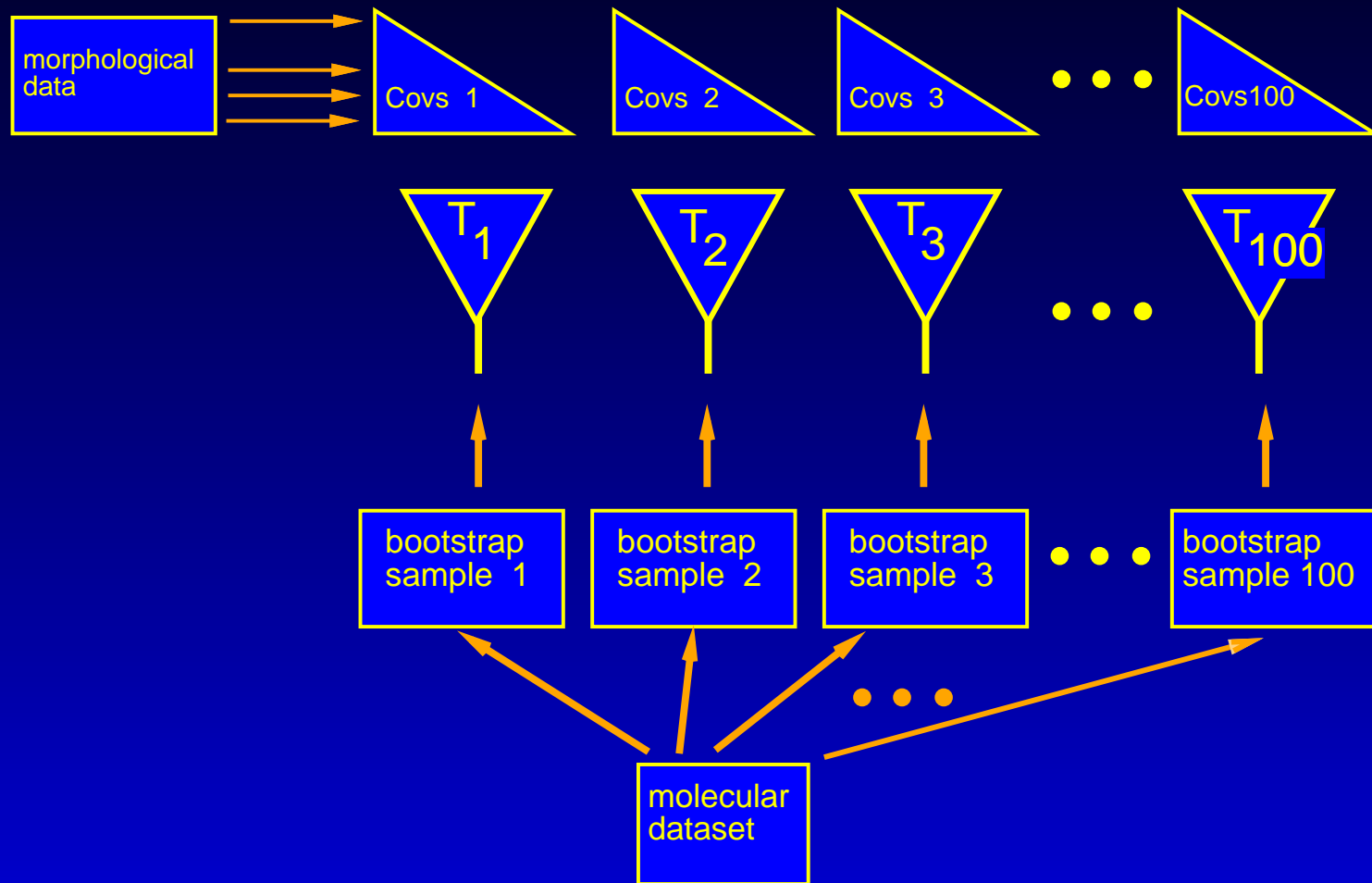
Propagating bootstrap sampling



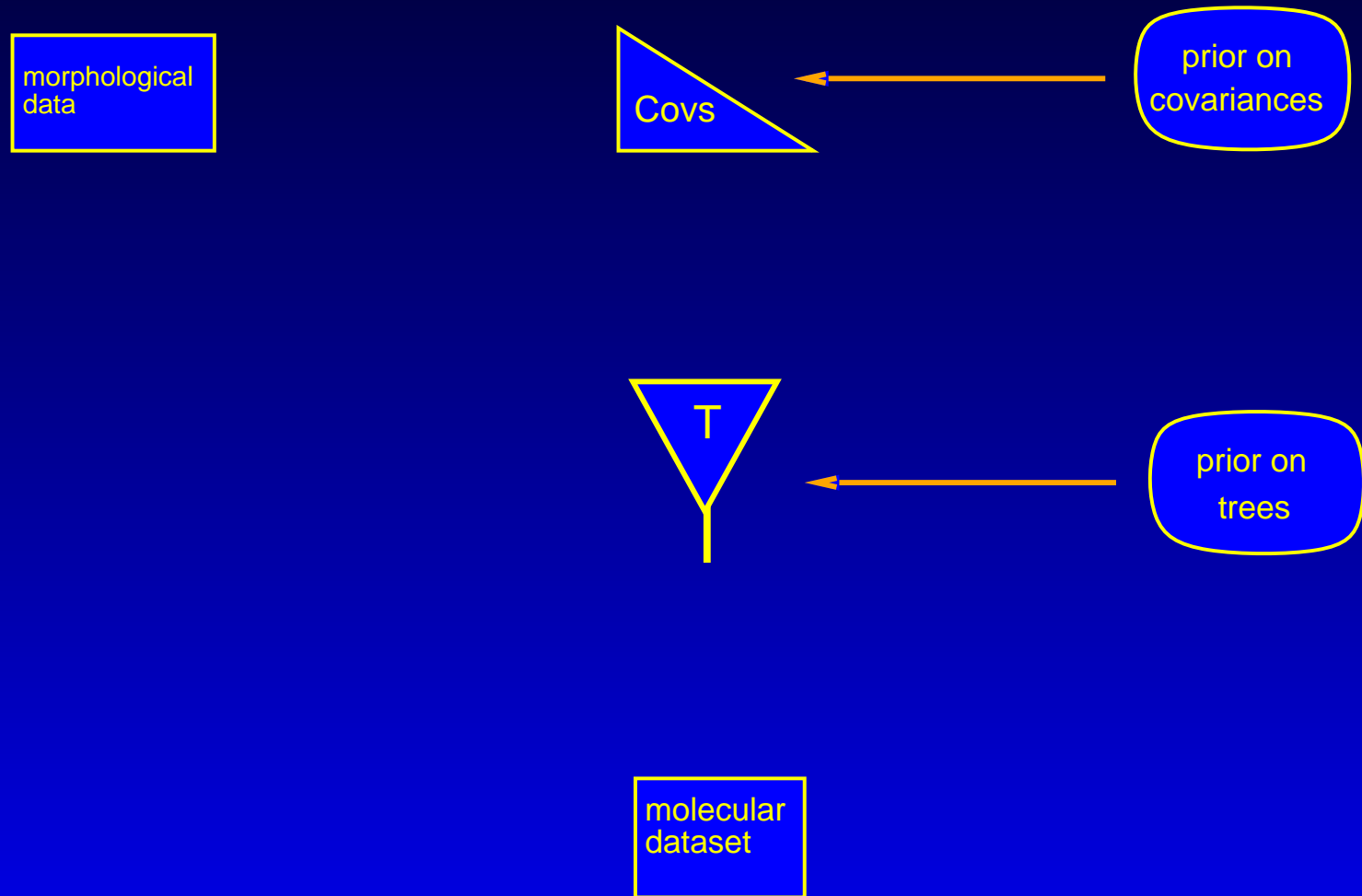
Propagating bootstrap sampling



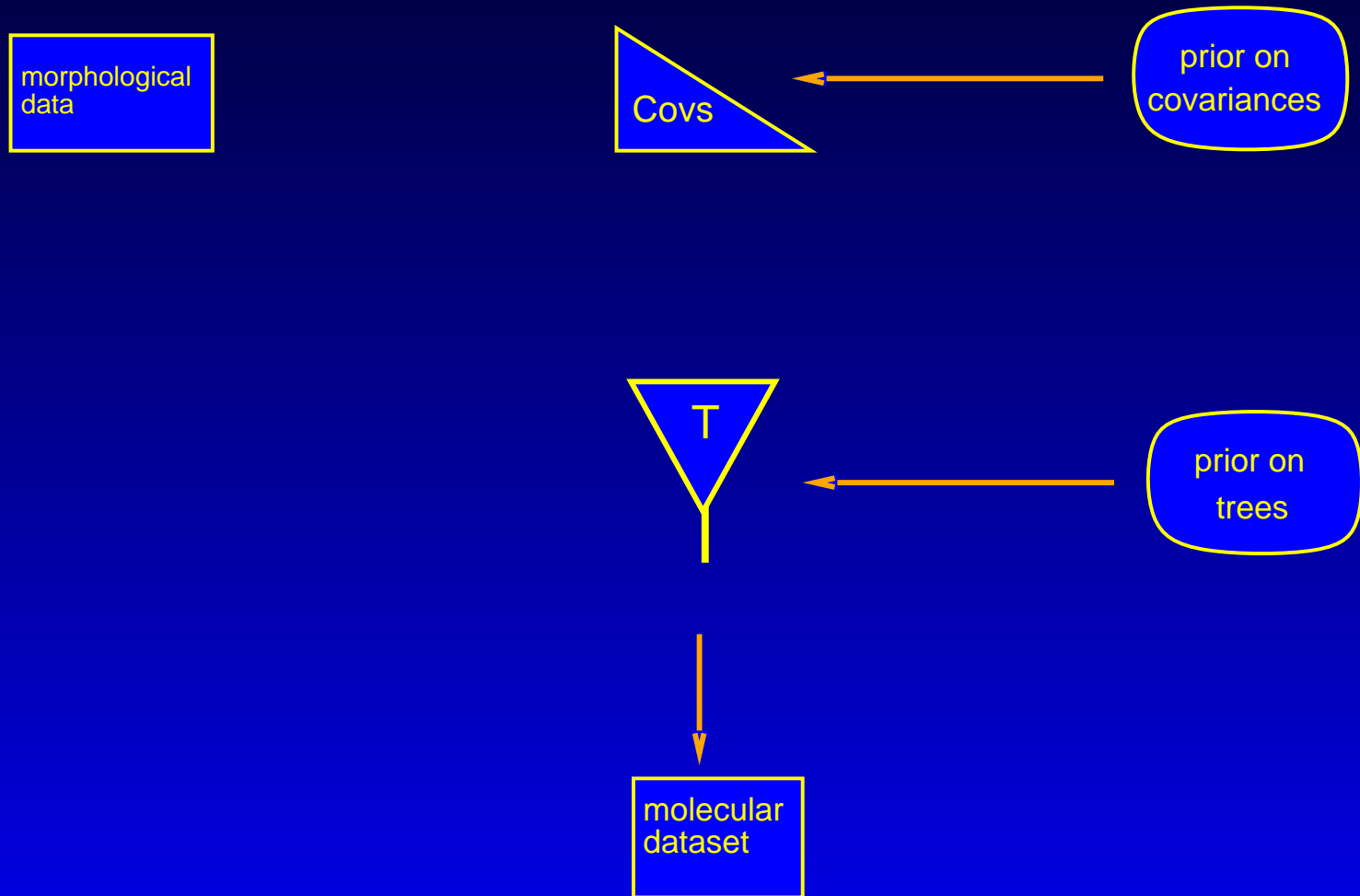
Propagating bootstrap sampling



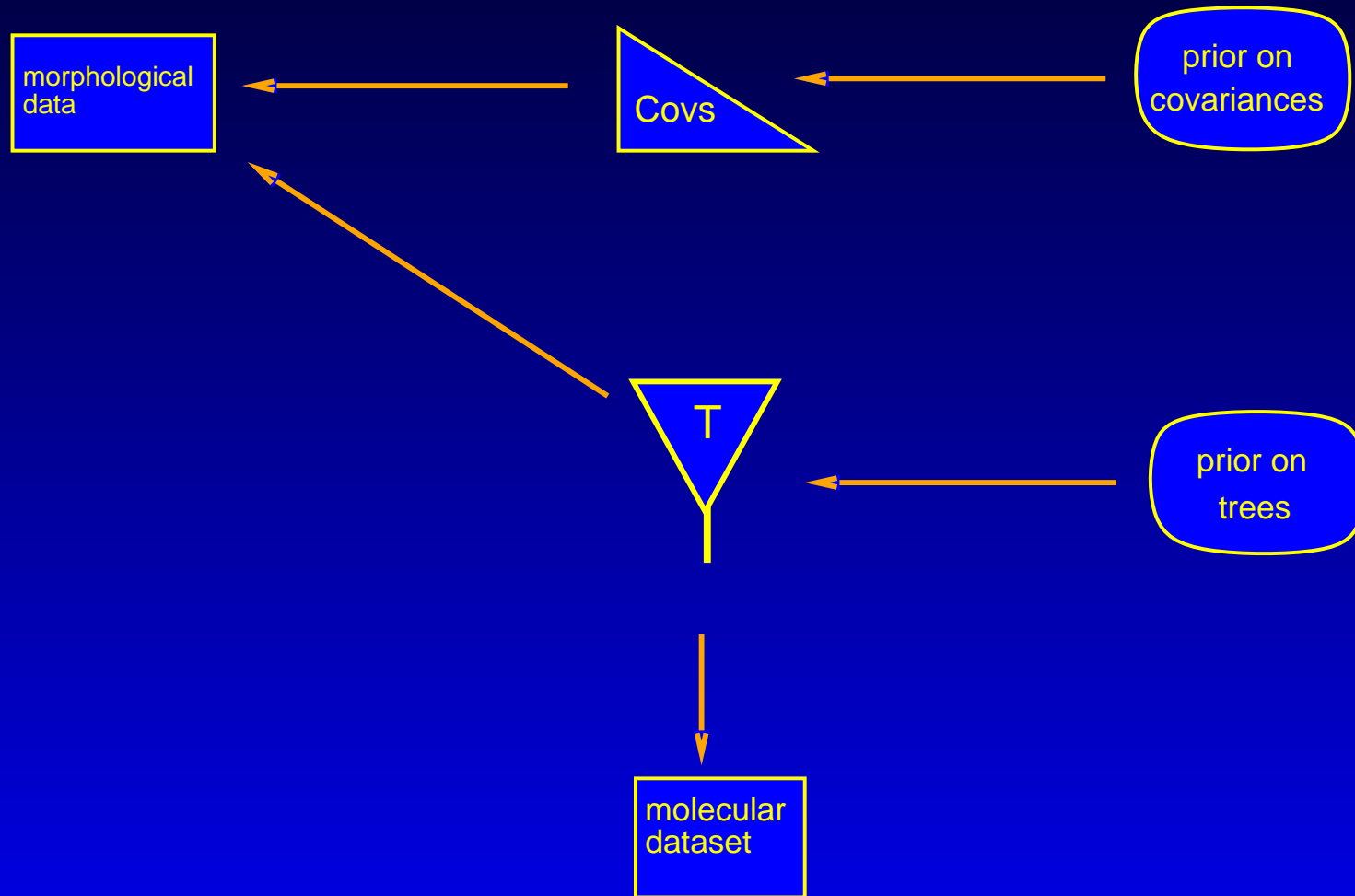
A Bayesian model



A Bayesian model

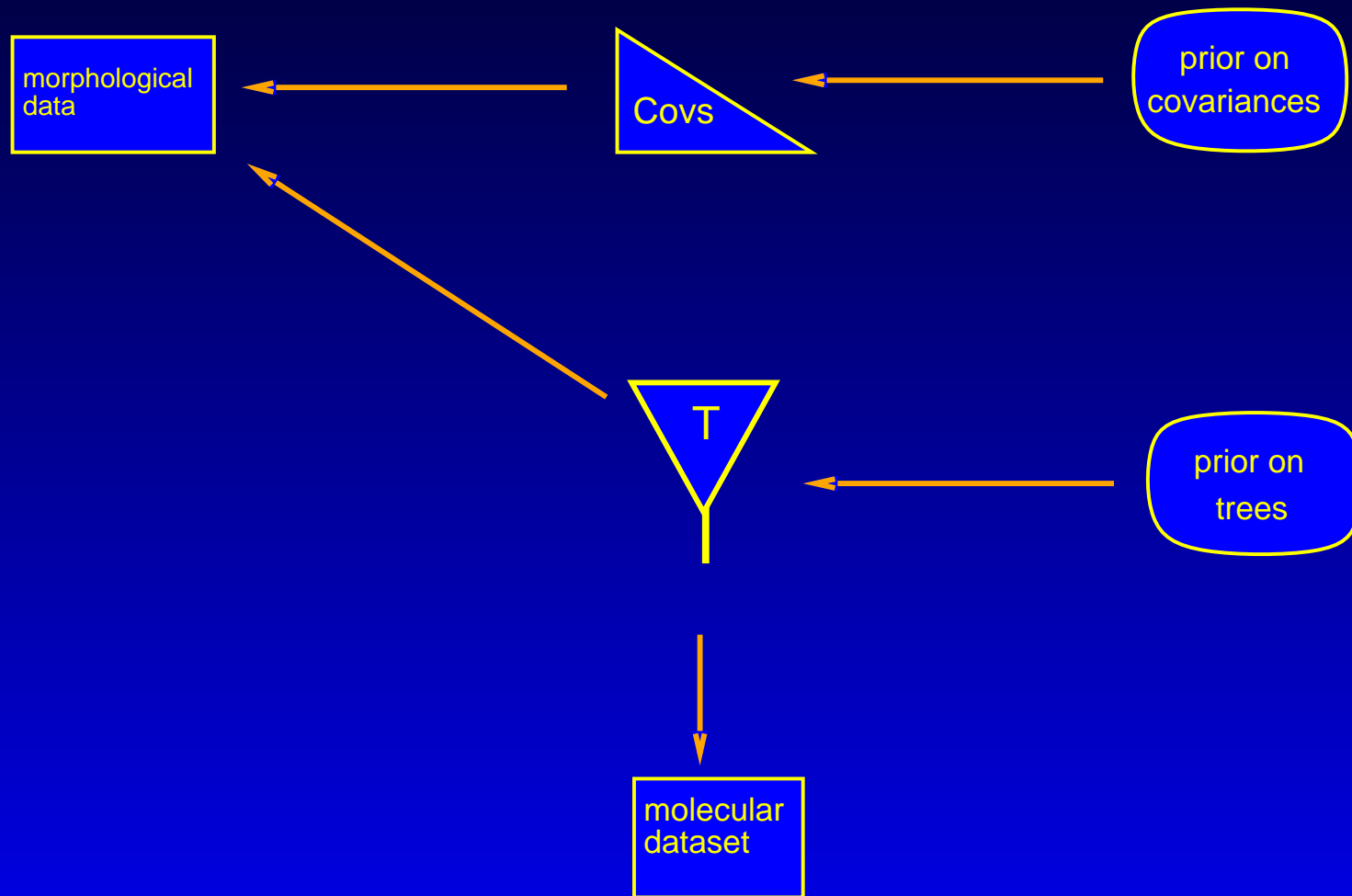


A Bayesian model

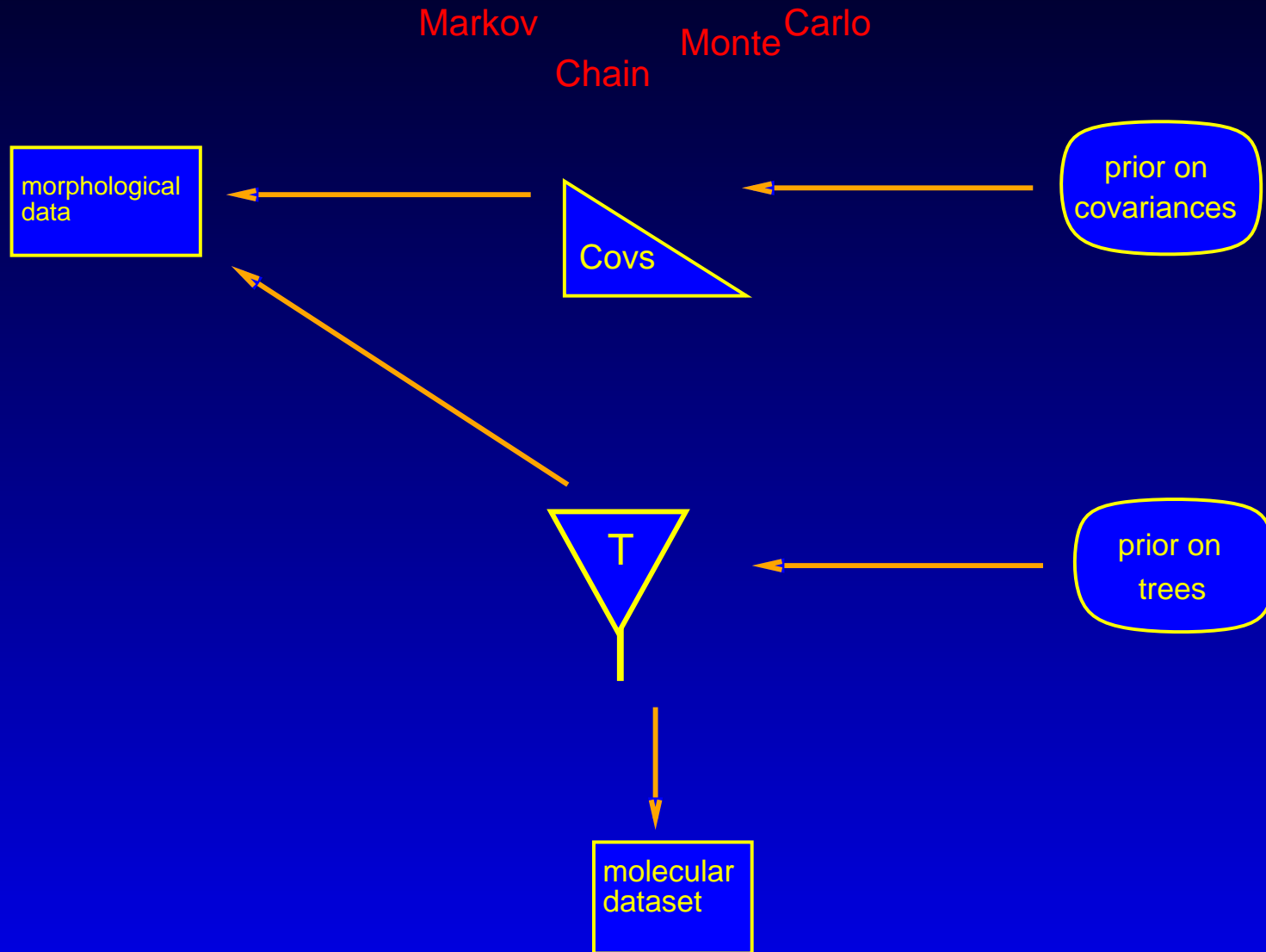


Bayesian MCMC

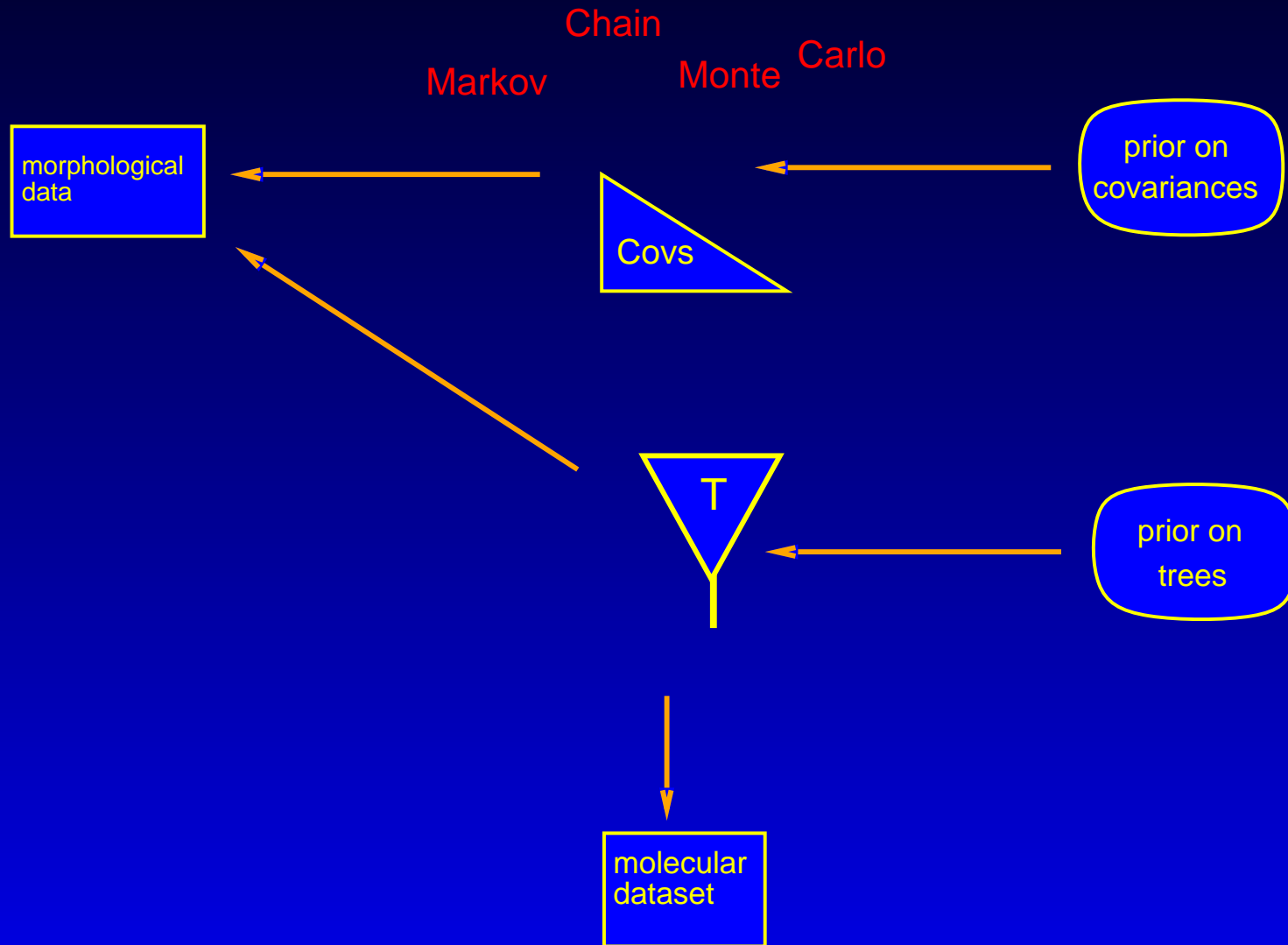
Markov Chain Monte Carlo



Bayesian MCMC

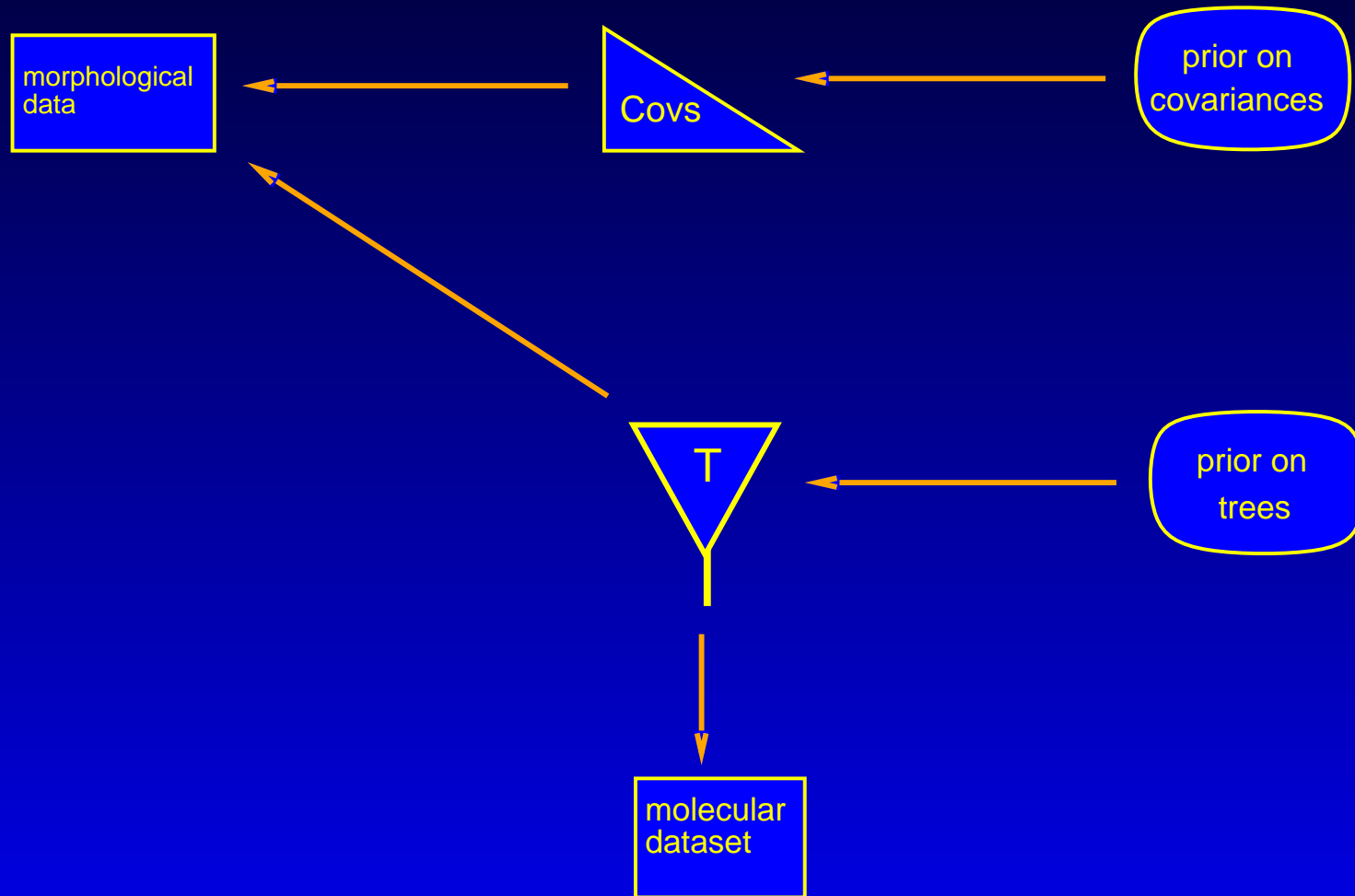


Bayesian MCMC

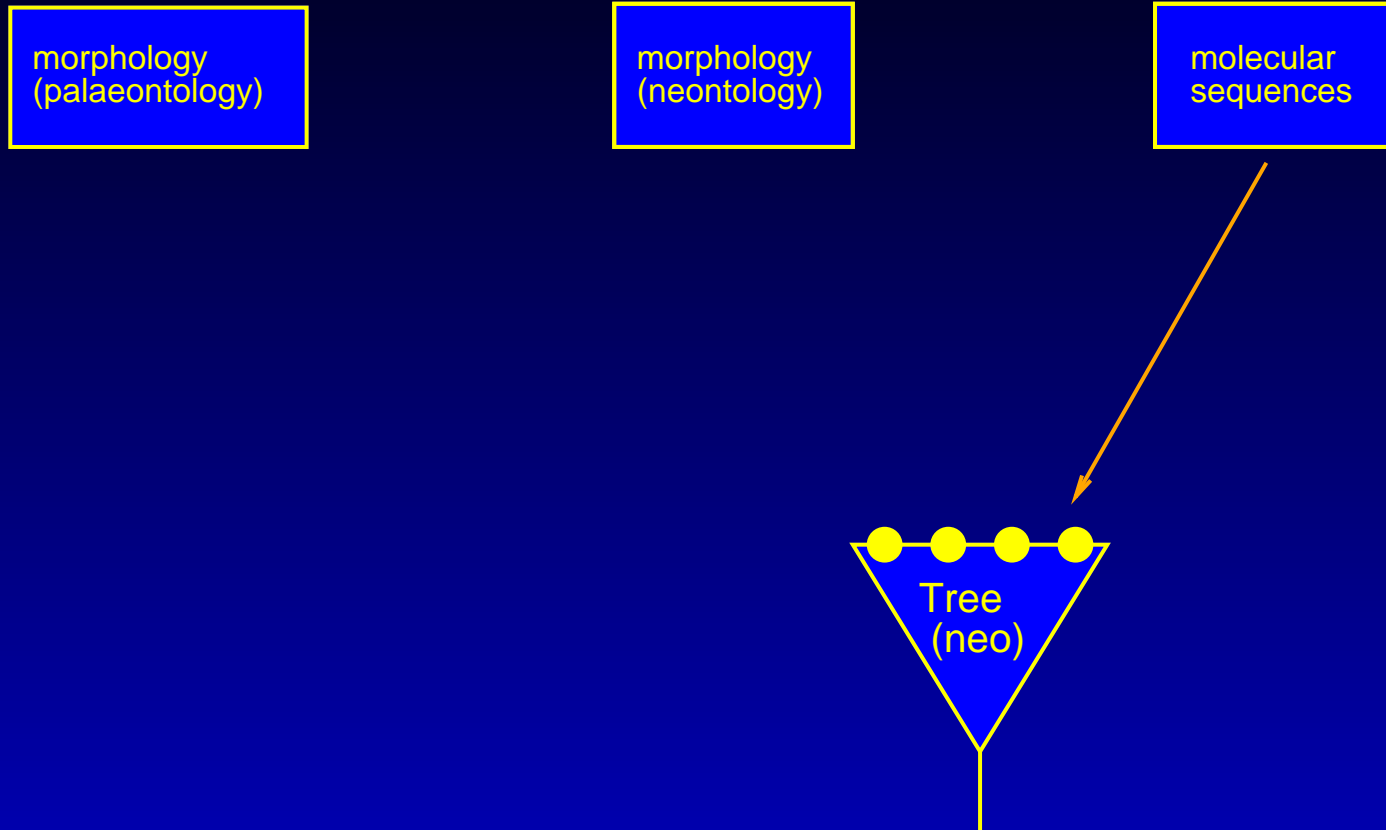


Bayesian MCMC

Markov Chain Monte Carlo

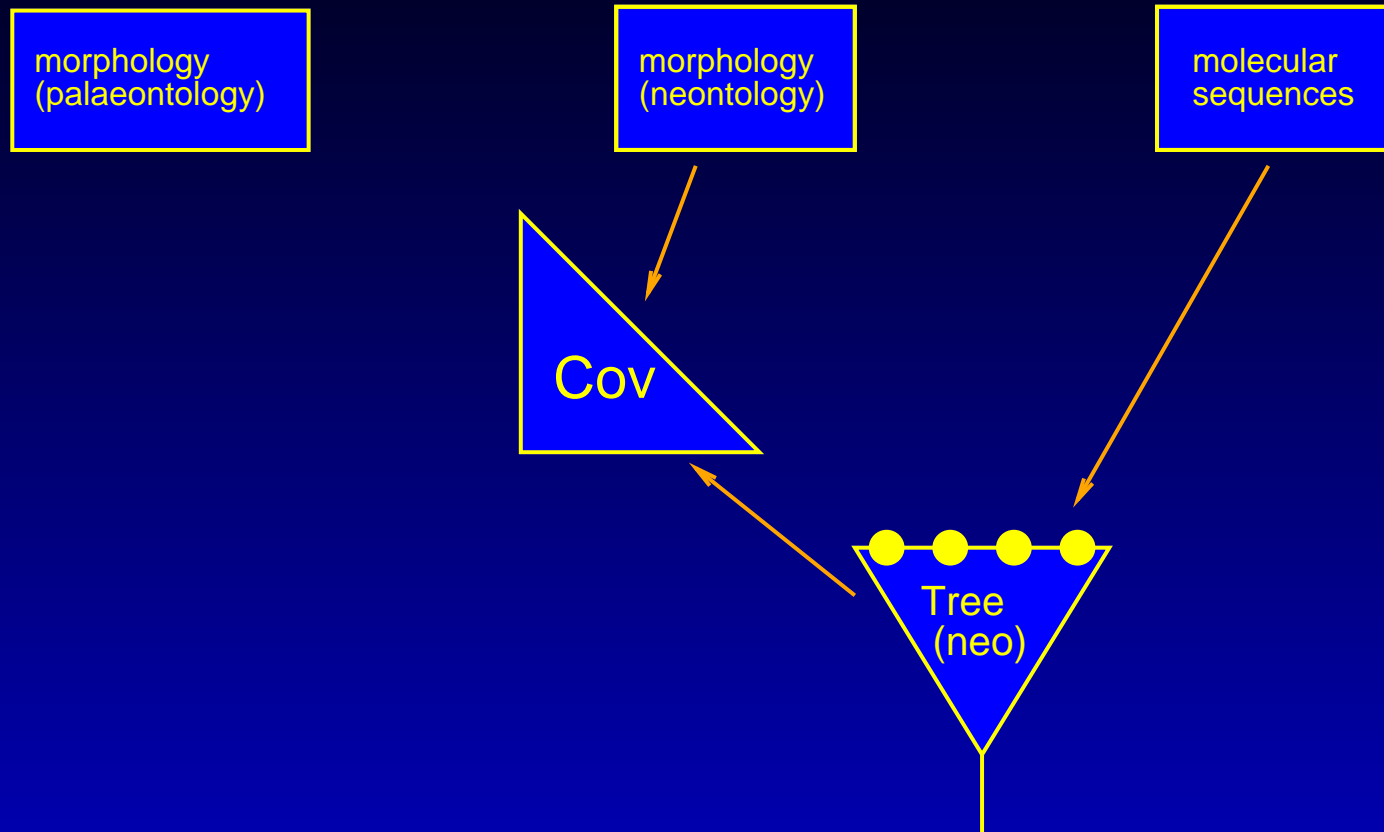


Using fossils sensibly



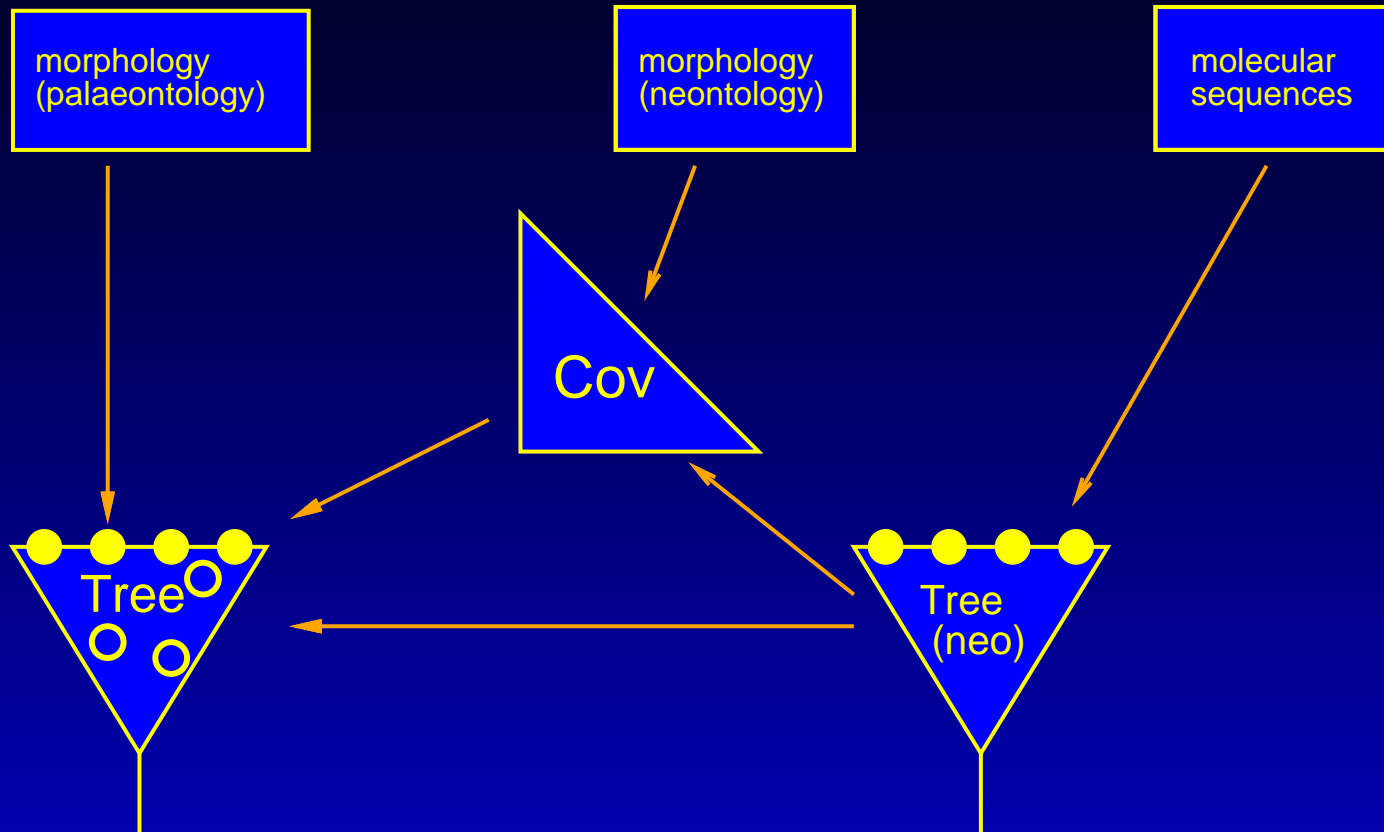
Infer tree of present-day species from molecular sequences

Using fossils sensibly



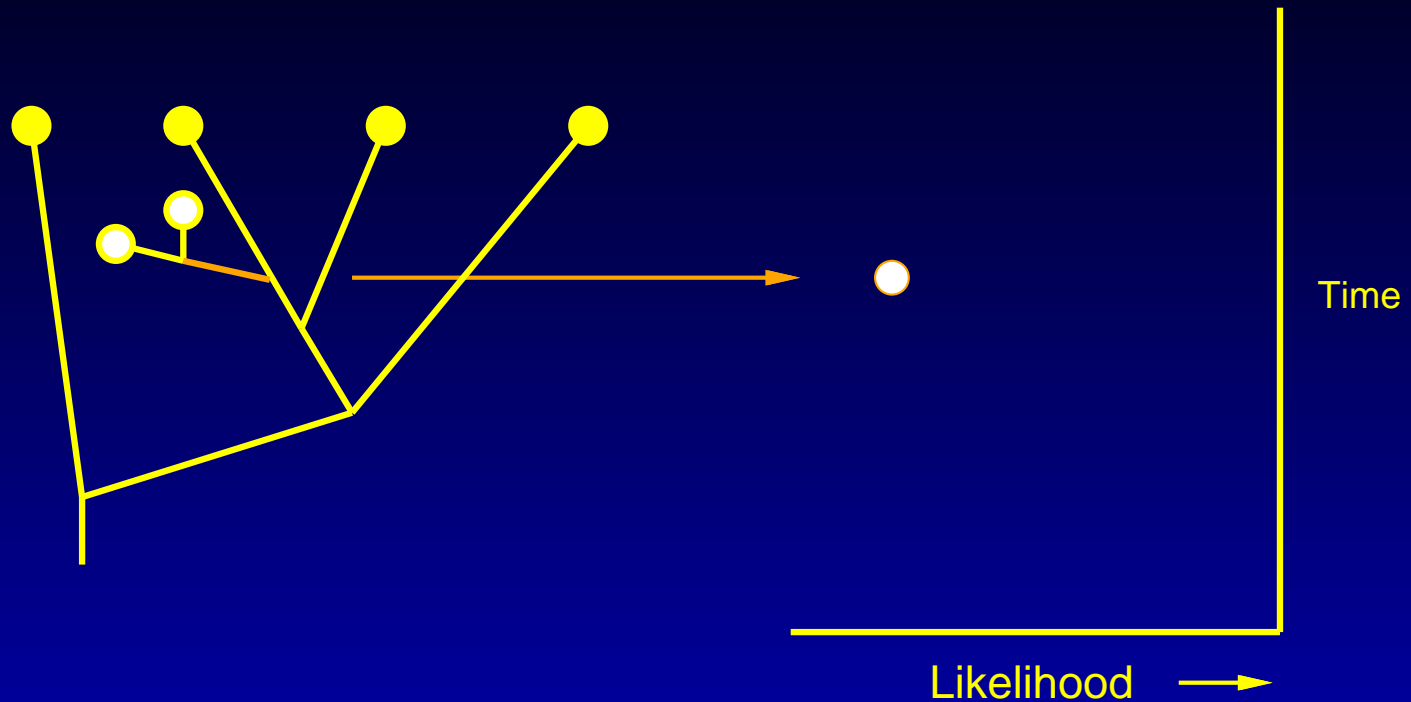
Infer covariances of morphology using it, present-day species

Using fossils sensibly



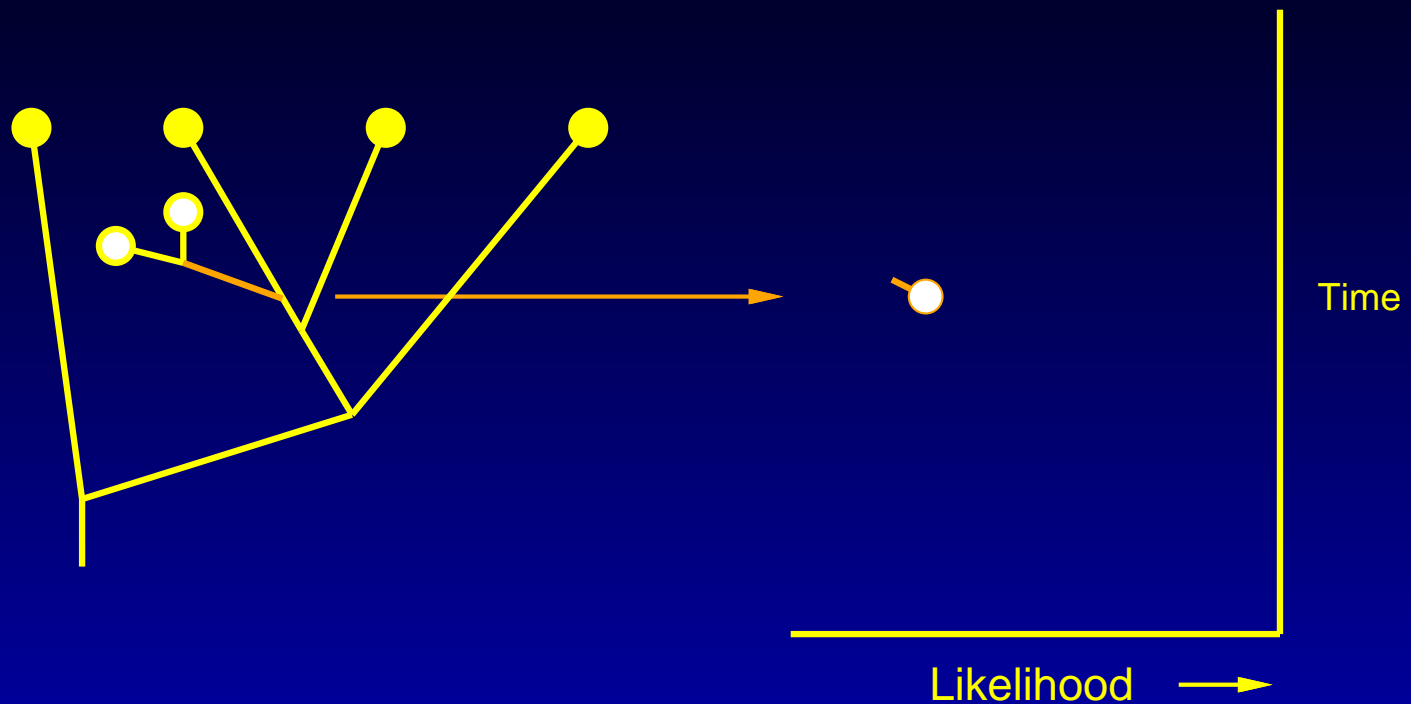
Infer placement of fossil species using their data

Using fossils sensibly



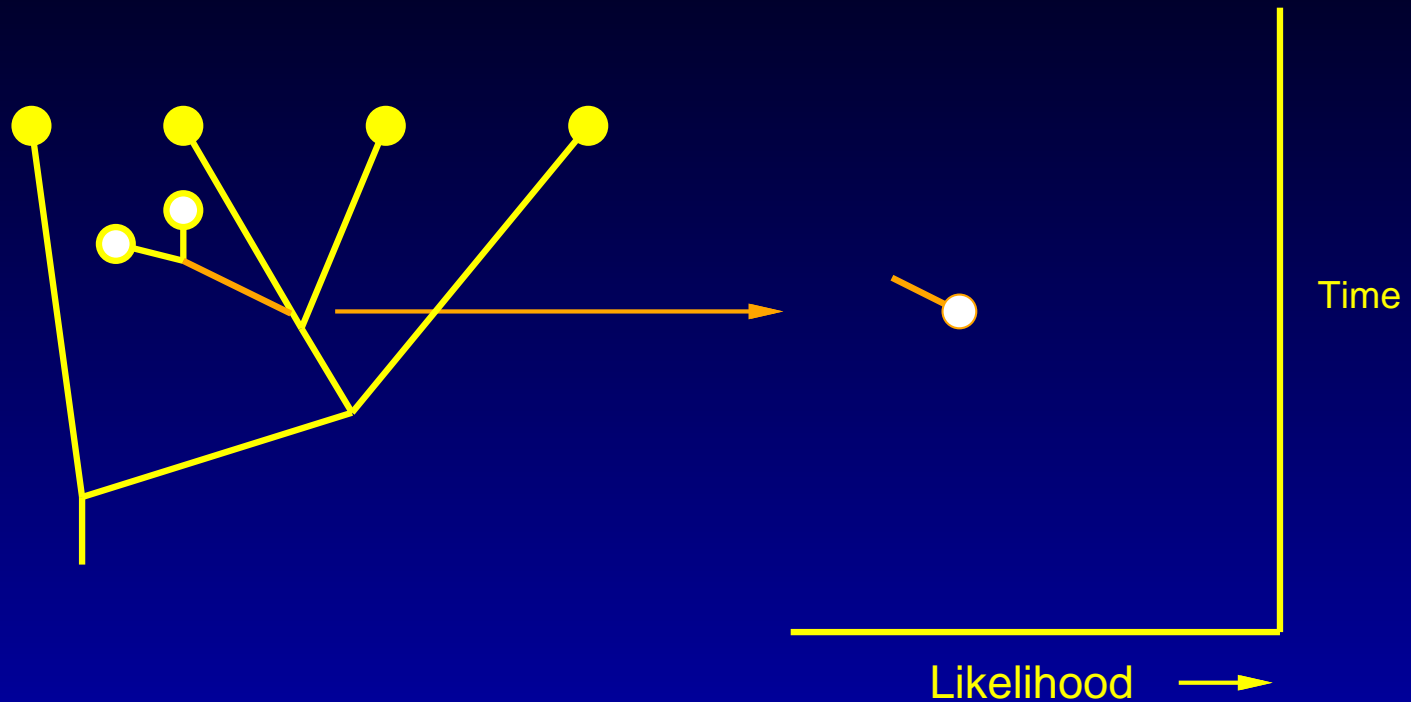
Use fossil and present-day morphology, covariances, tree, also stratigraphic models

Using fossils sensibly



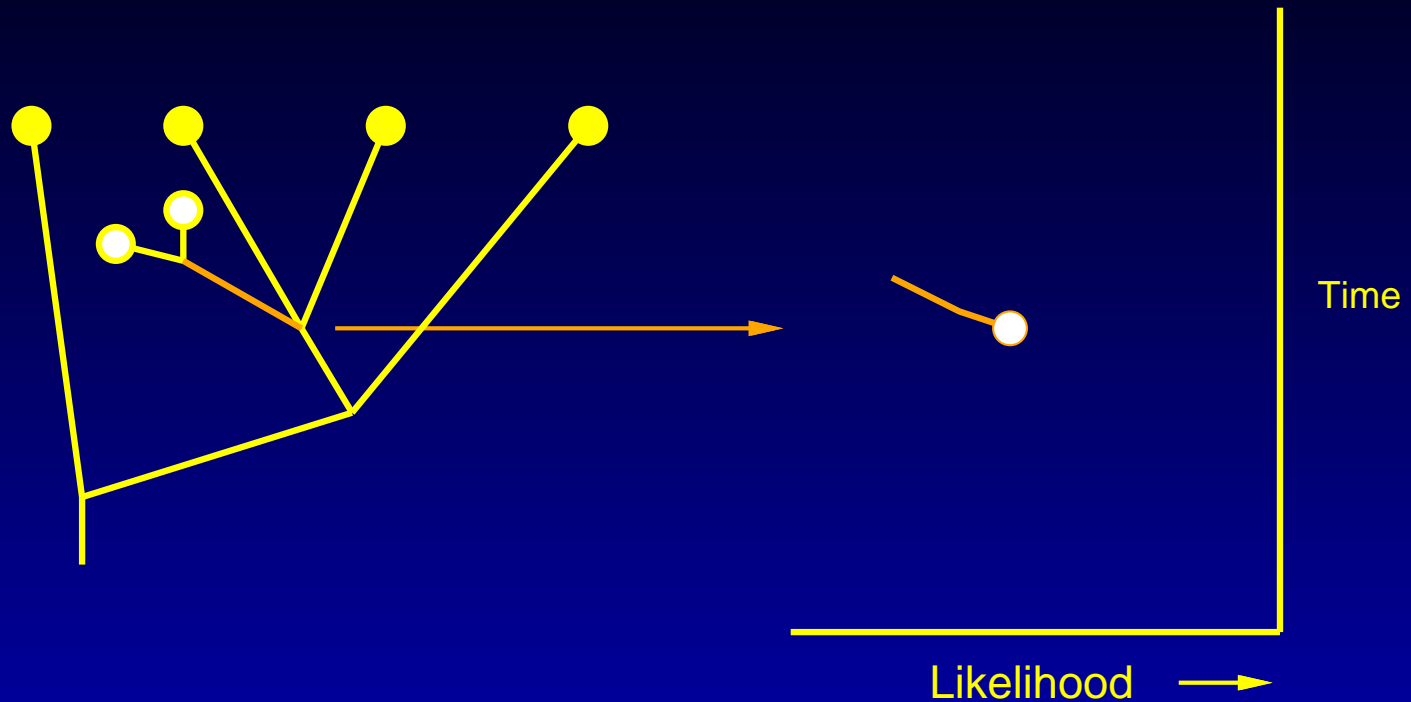
Use fossil and present-day morphology, covariances, tree, also stratigraphic models

Using fossils sensibly



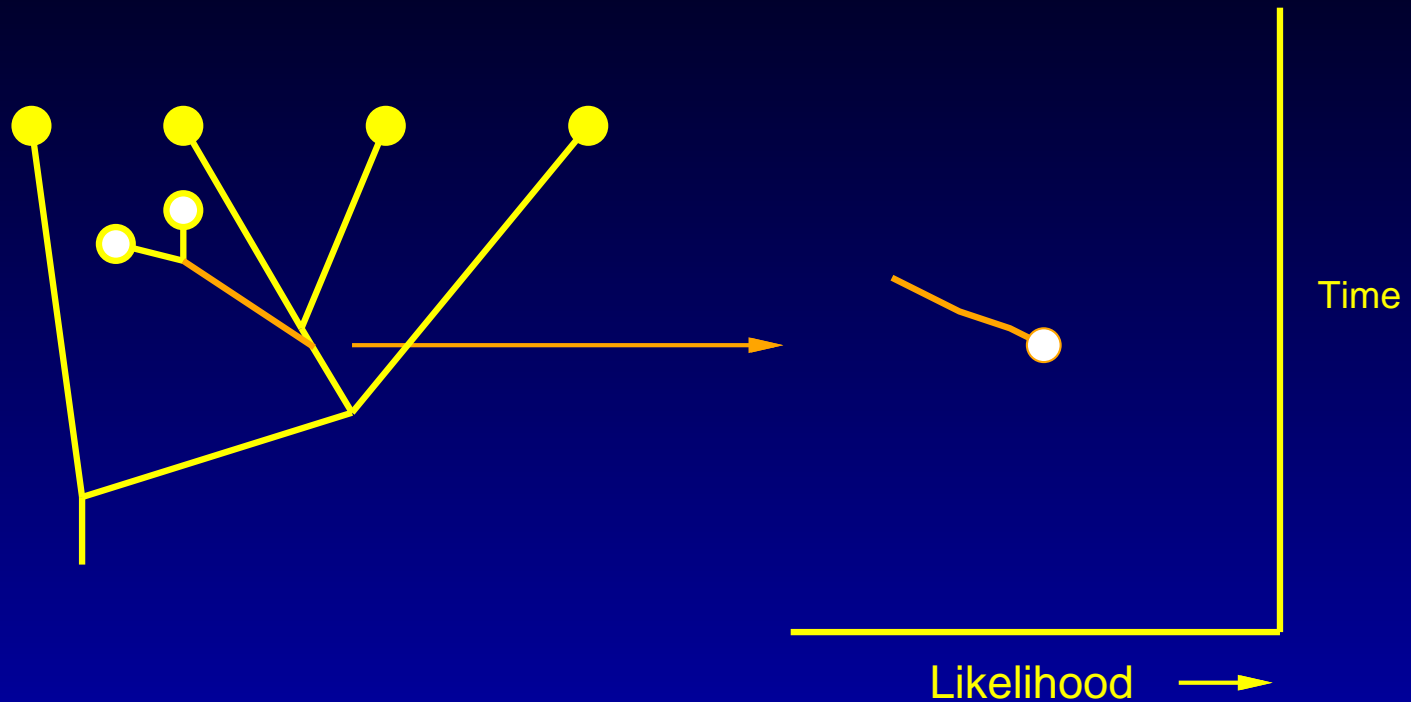
Use fossil and present-day morphology, covariances, tree, also stratigraphic models

Using fossils sensibly



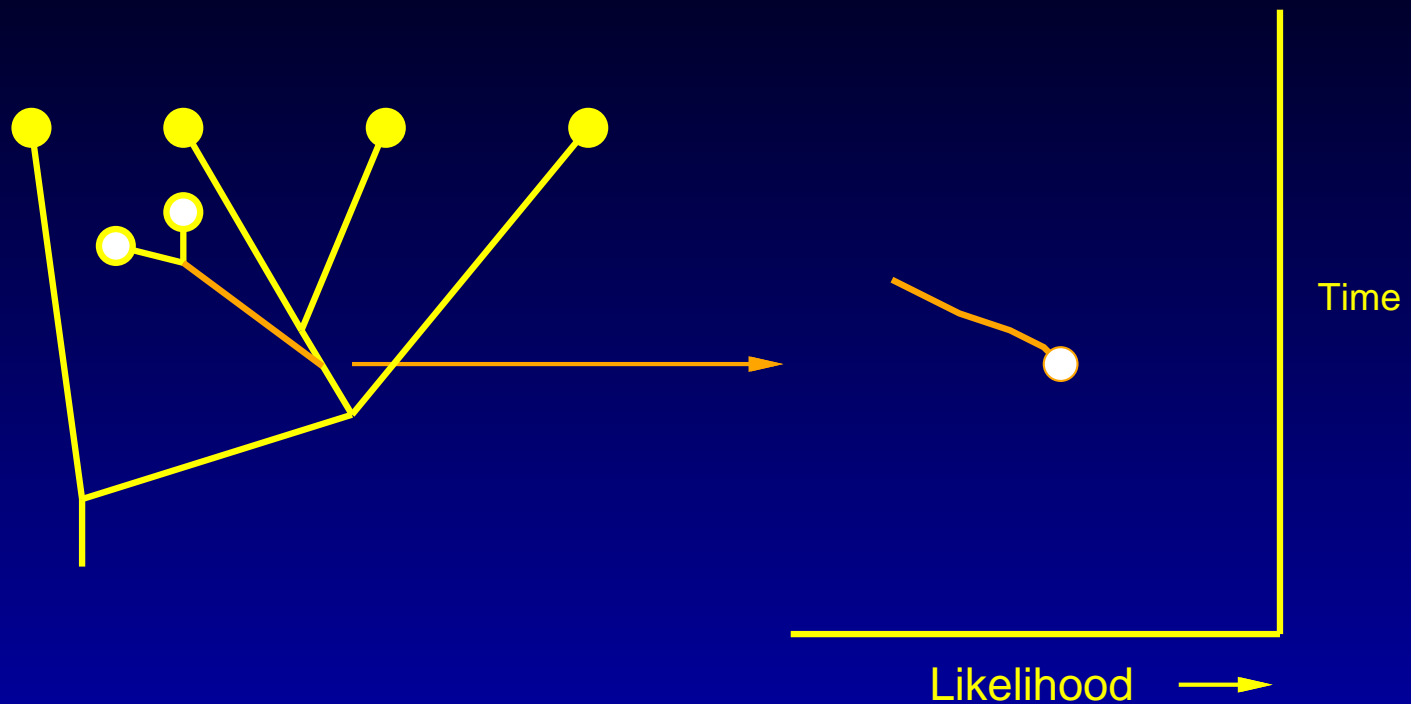
Use fossil and present-day morphology, covariances, tree, also stratigraphic models

Using fossils sensibly



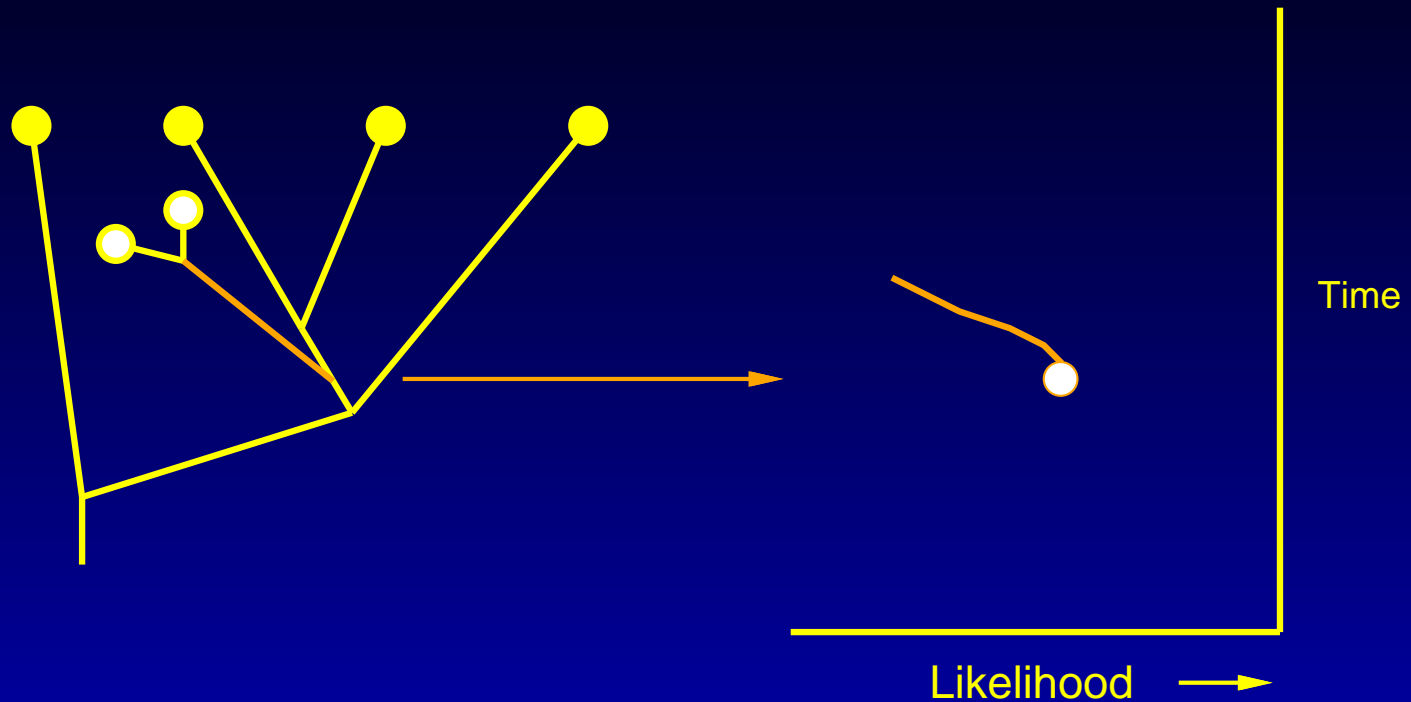
Use fossil and present-day morphology, covariances, tree, also stratigraphic models

Using fossils sensibly



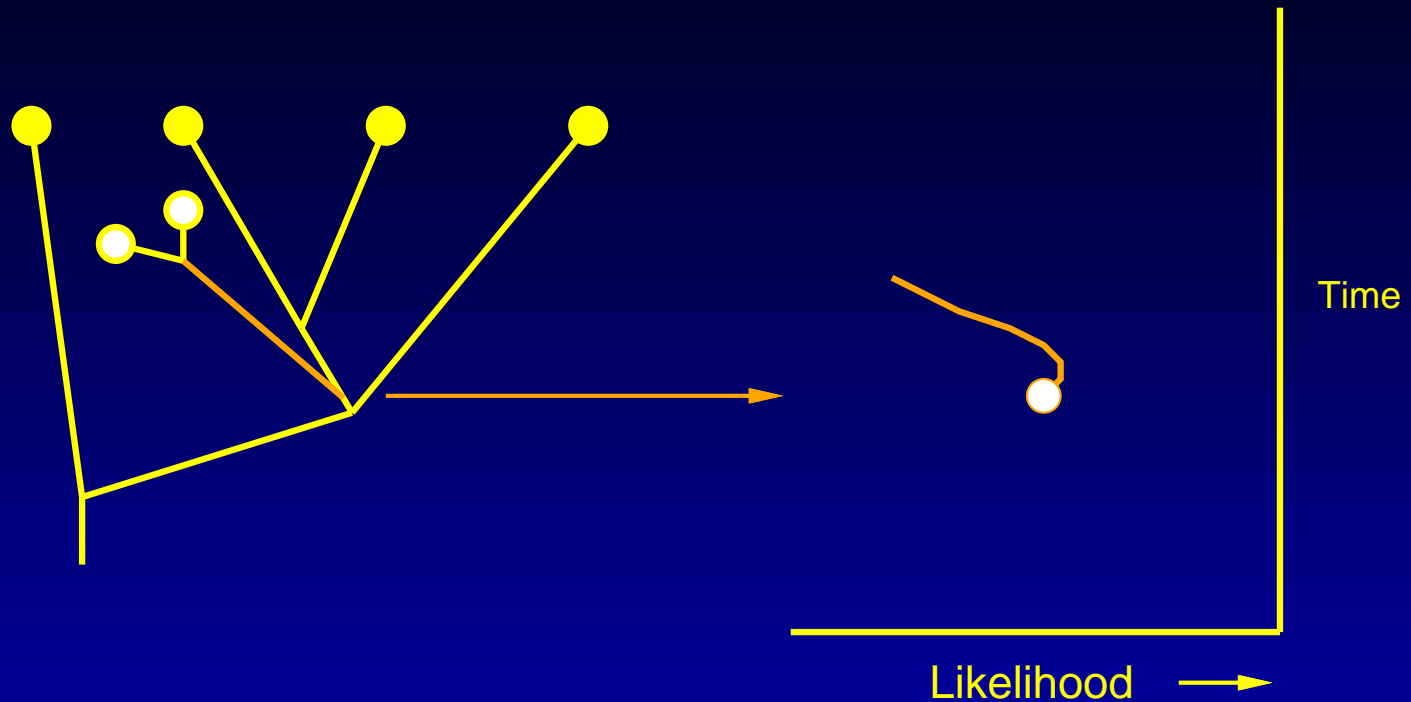
Use fossil and present-day morphology, covariances, tree, also stratigraphic models

Using fossils sensibly



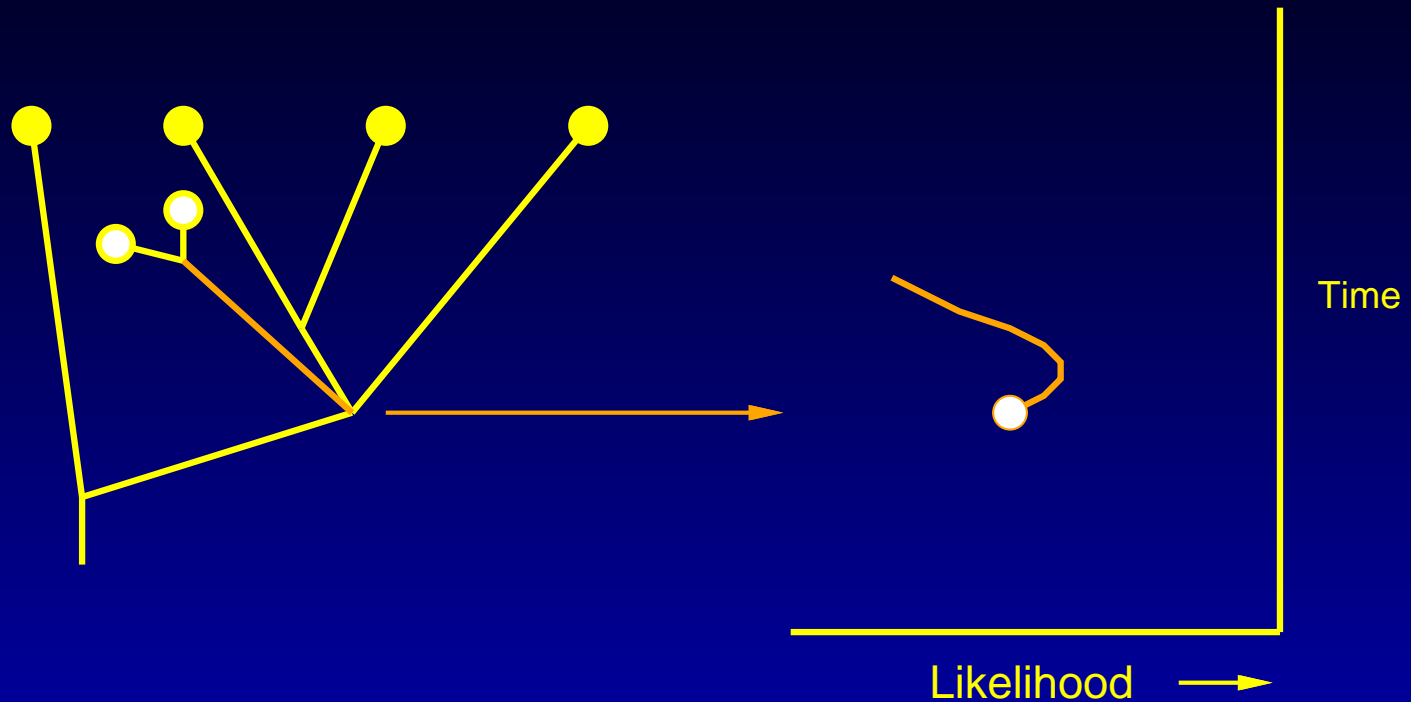
Use fossil and present-day morphology, covariances, tree, also stratigraphic models

Using fossils sensibly



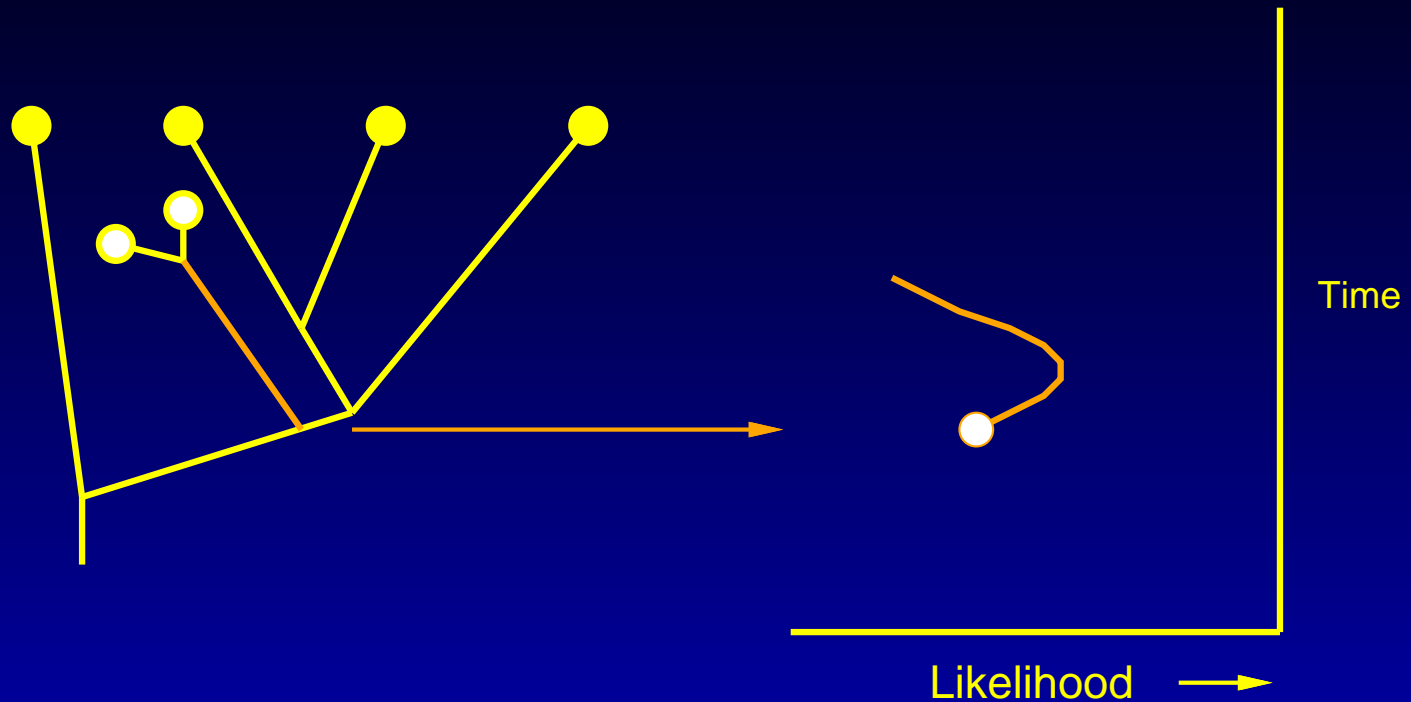
Use fossil and present-day morphology, covariances, tree,
also stratigraphic models

Using fossils sensibly



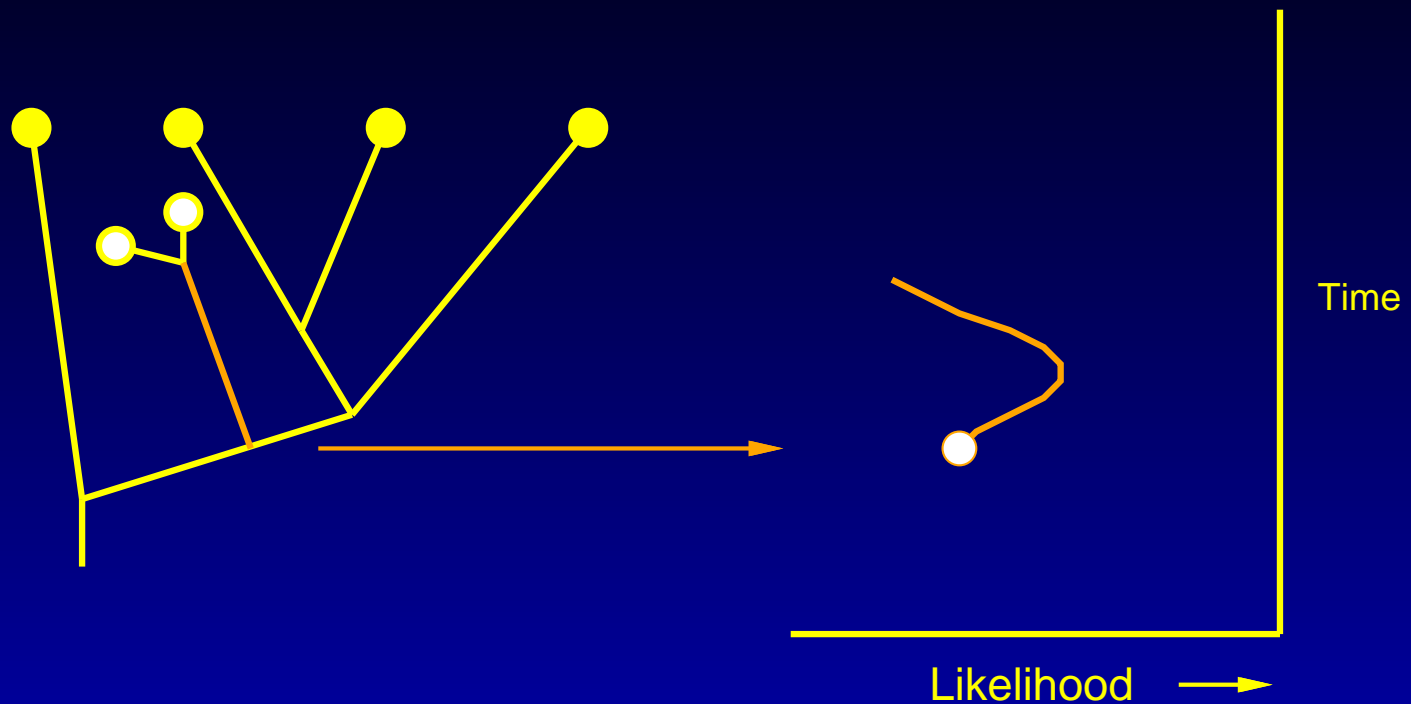
Use fossil and present-day morphology, covariances, tree, also stratigraphic models

Using fossils sensibly



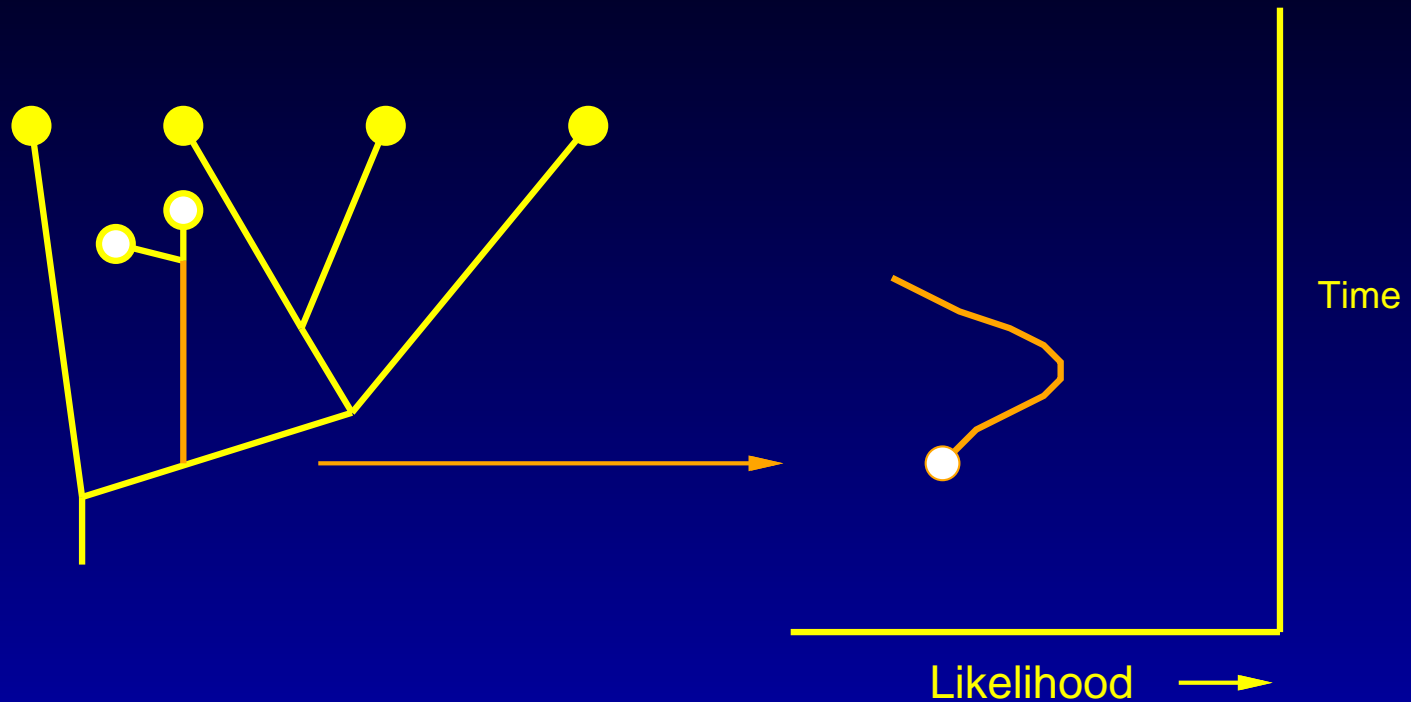
Use fossil and present-day morphology, covariances, tree, also stratigraphic models

Using fossils sensibly



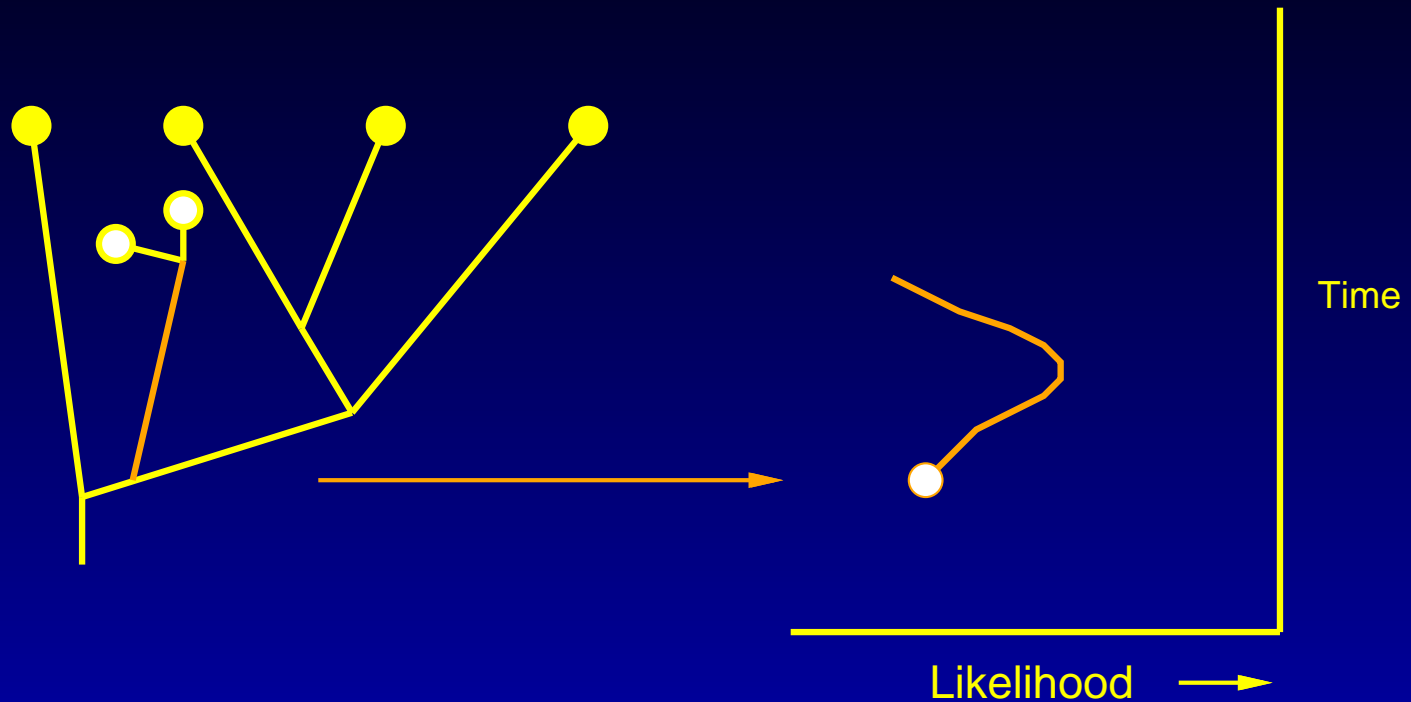
Use fossil and present-day morphology, covariances, tree,
also stratigraphic models

Using fossils sensibly



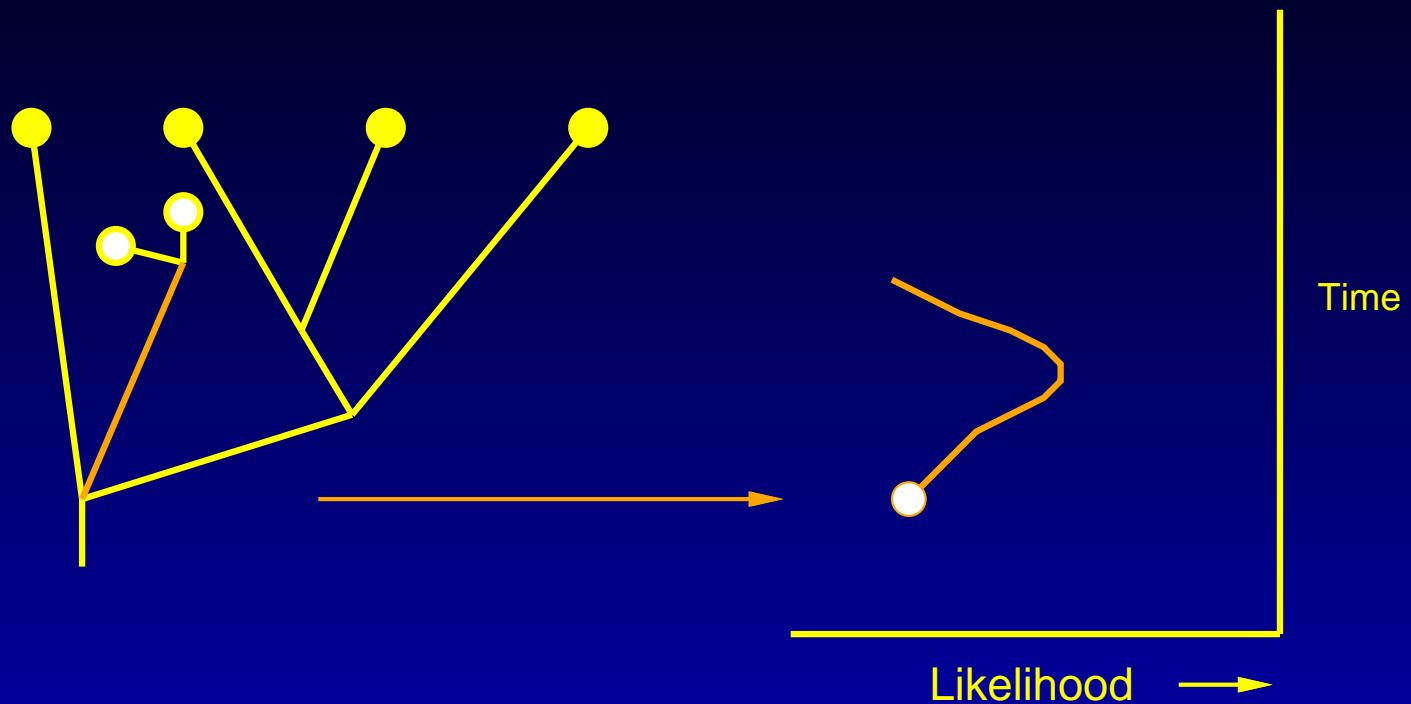
Use fossil and present-day morphology, covariances, tree, also stratigraphic models

Using fossils sensibly



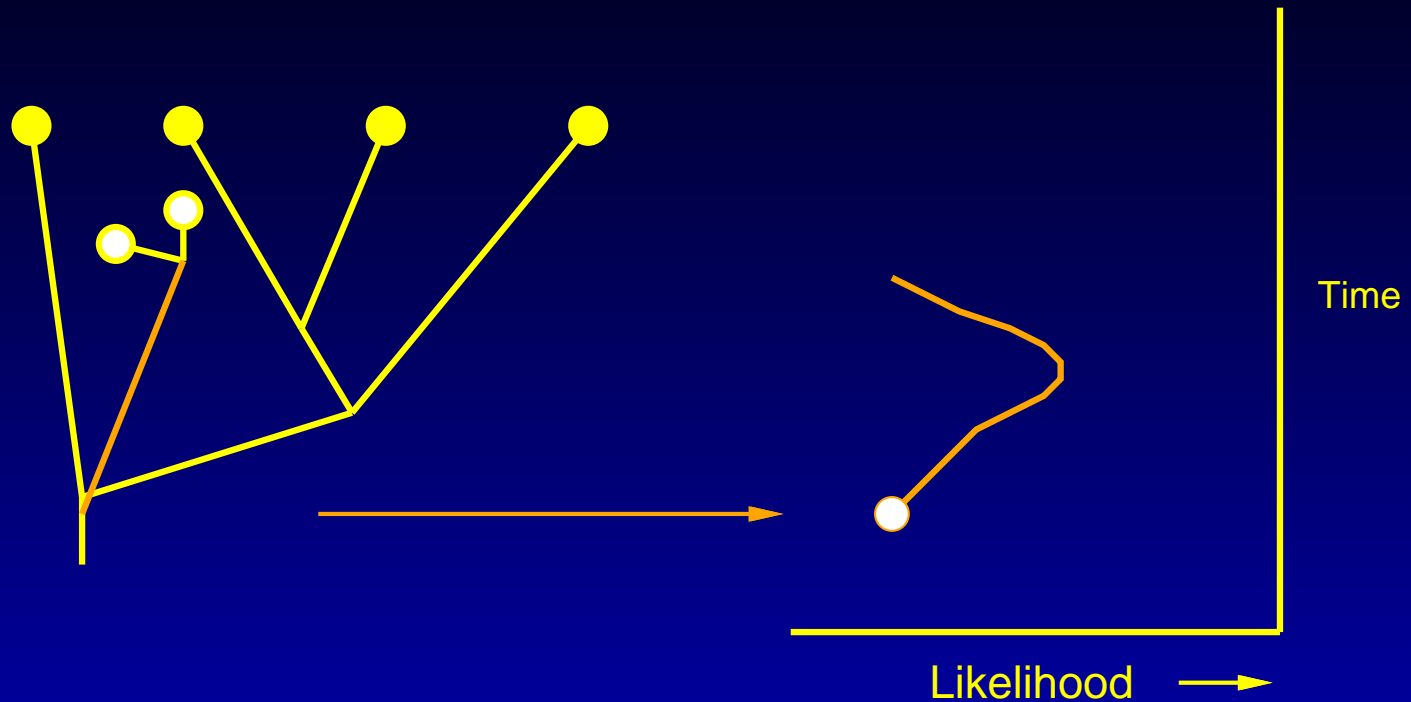
Use fossil and present-day morphology, covariances, tree, also stratigraphic models

Using fossils sensibly



Use fossil and present-day morphology, covariances, tree, also stratigraphic models

Using fossils sensibly



Use fossil and present-day morphology, covariances, tree, also stratigraphic models

Physical Anthropology – current methodology?

(An outsider's caricature)

- Find one bone

Physical Anthropology – current methodology?

(An outsider's caricature)

- Find one bone
- Hold a press conference

Physical Anthropology – current methodology?

(An outsider's caricature)

- Find one bone
- Hold a press conference
- Announce that this is a new genus

Physical Anthropology – current methodology?

(An outsider's caricature)

- Find one bone
- Hold a press conference
- Announce that this is a new genus
- Announce that it finally solves all the problems of human evolution

Physical Anthropology – current methodology?

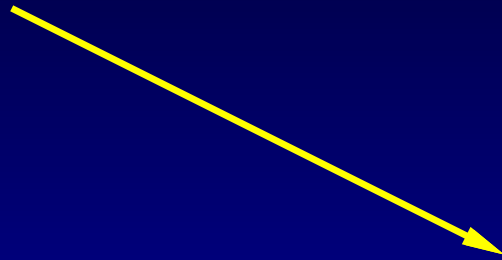
(An outsider's caricature)

- Find one bone
- Hold a press conference
- Announce that this is a new genus
- Announce that it finally solves all the problems of human evolution

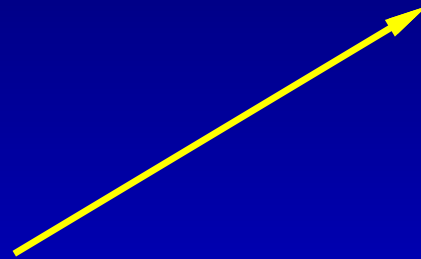
This creates pressure for splitting species, ignoring within-species variation, and overinterpreting data.

Then: the importance of (beta) classification!

Phenotypic Data

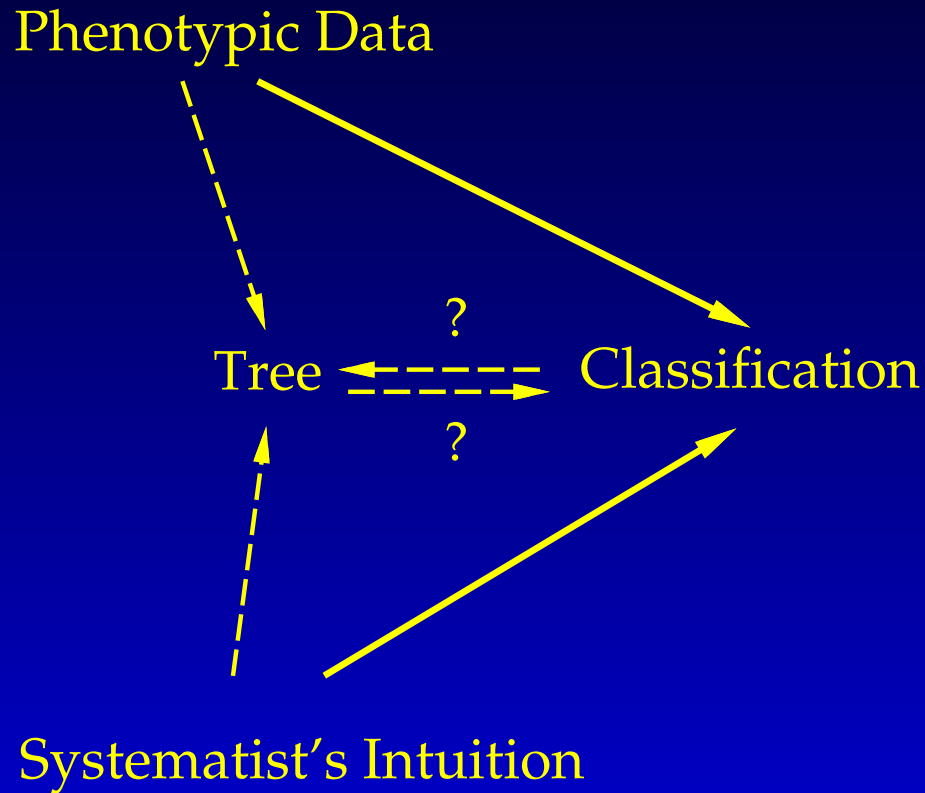


Classification

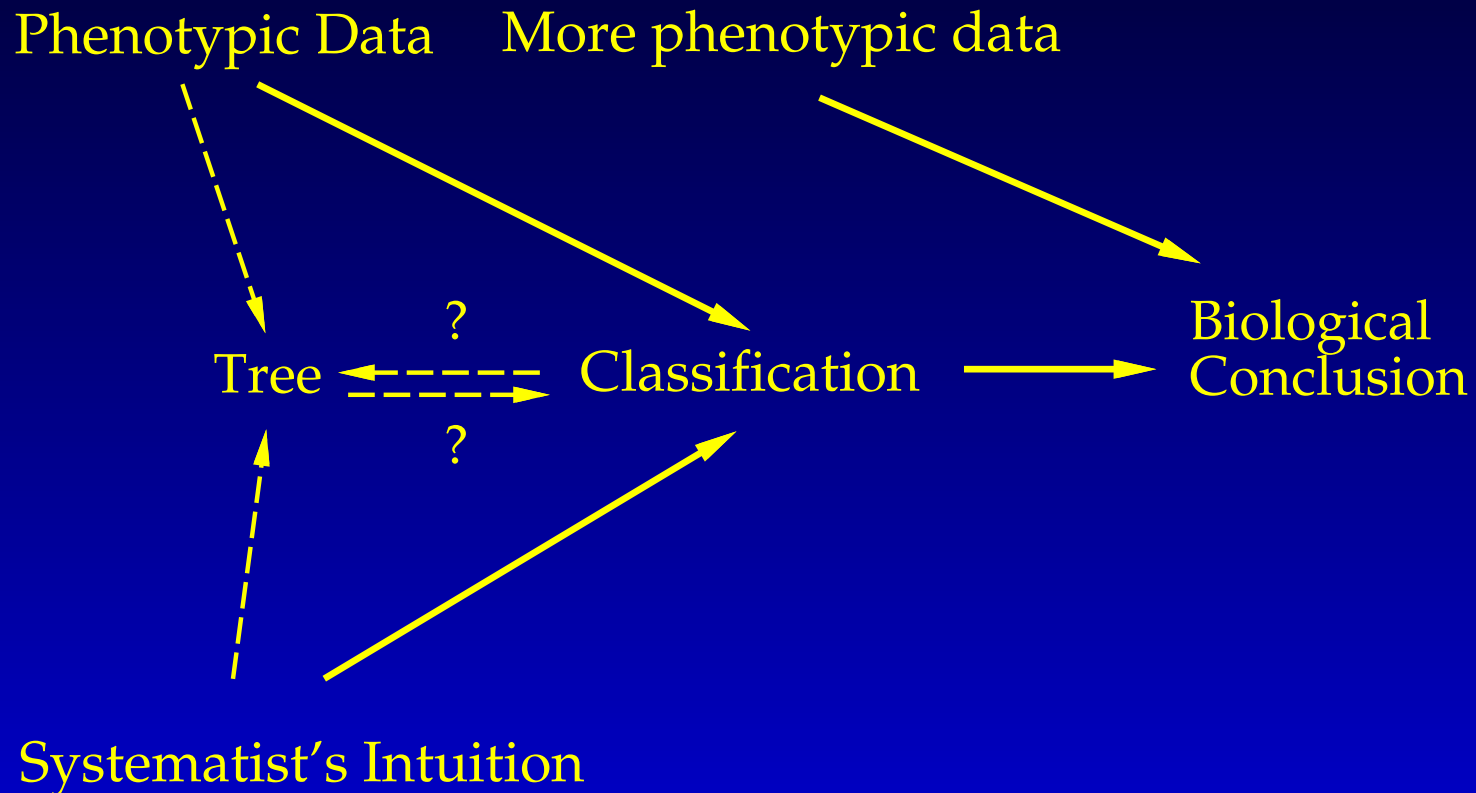


Systematist's Intuition

Then: the importance of (beta) classification!



Then: the importance of (beta) classification!



Now: the importance of (beta) classification?

Molecular Data

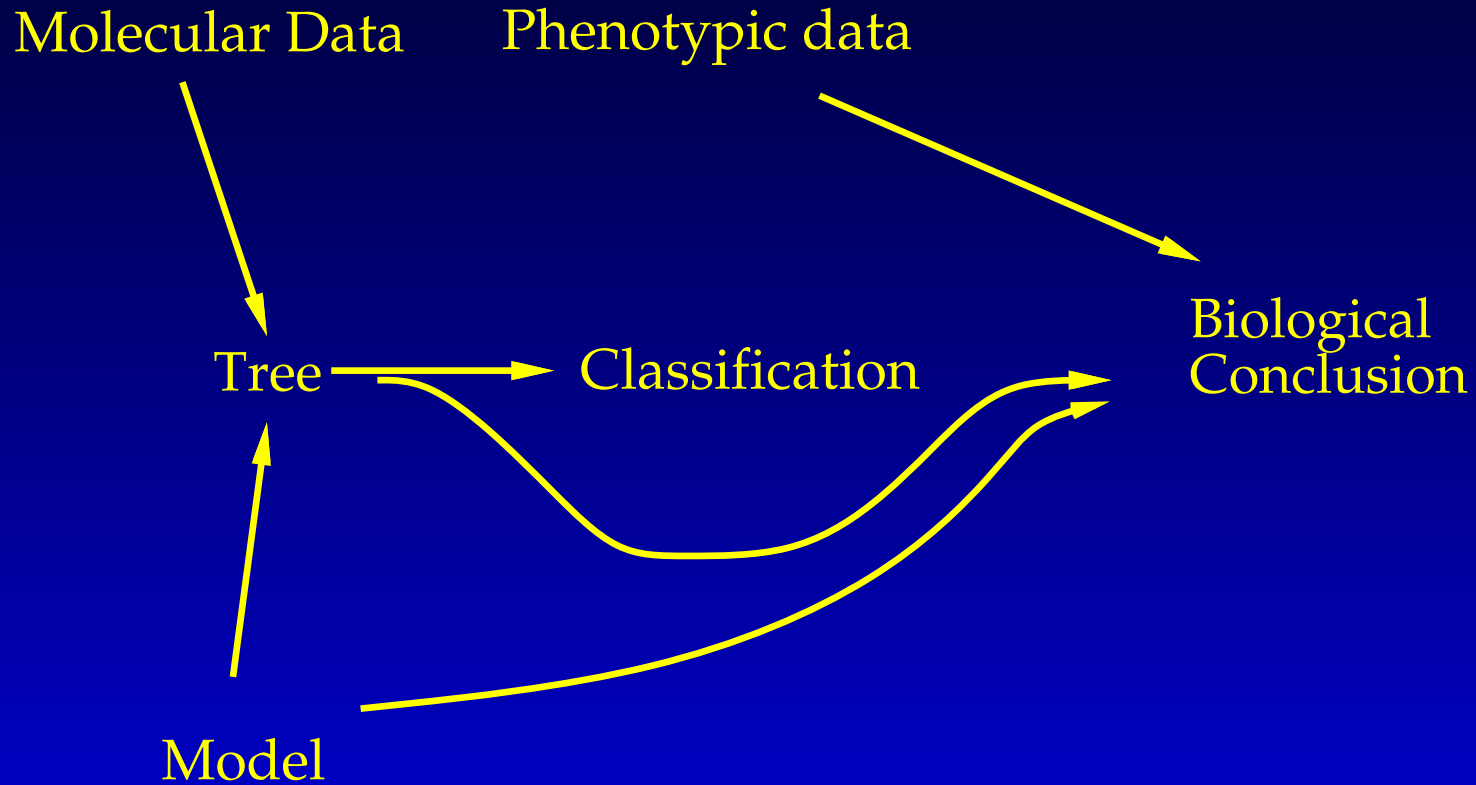


Tree → Classification

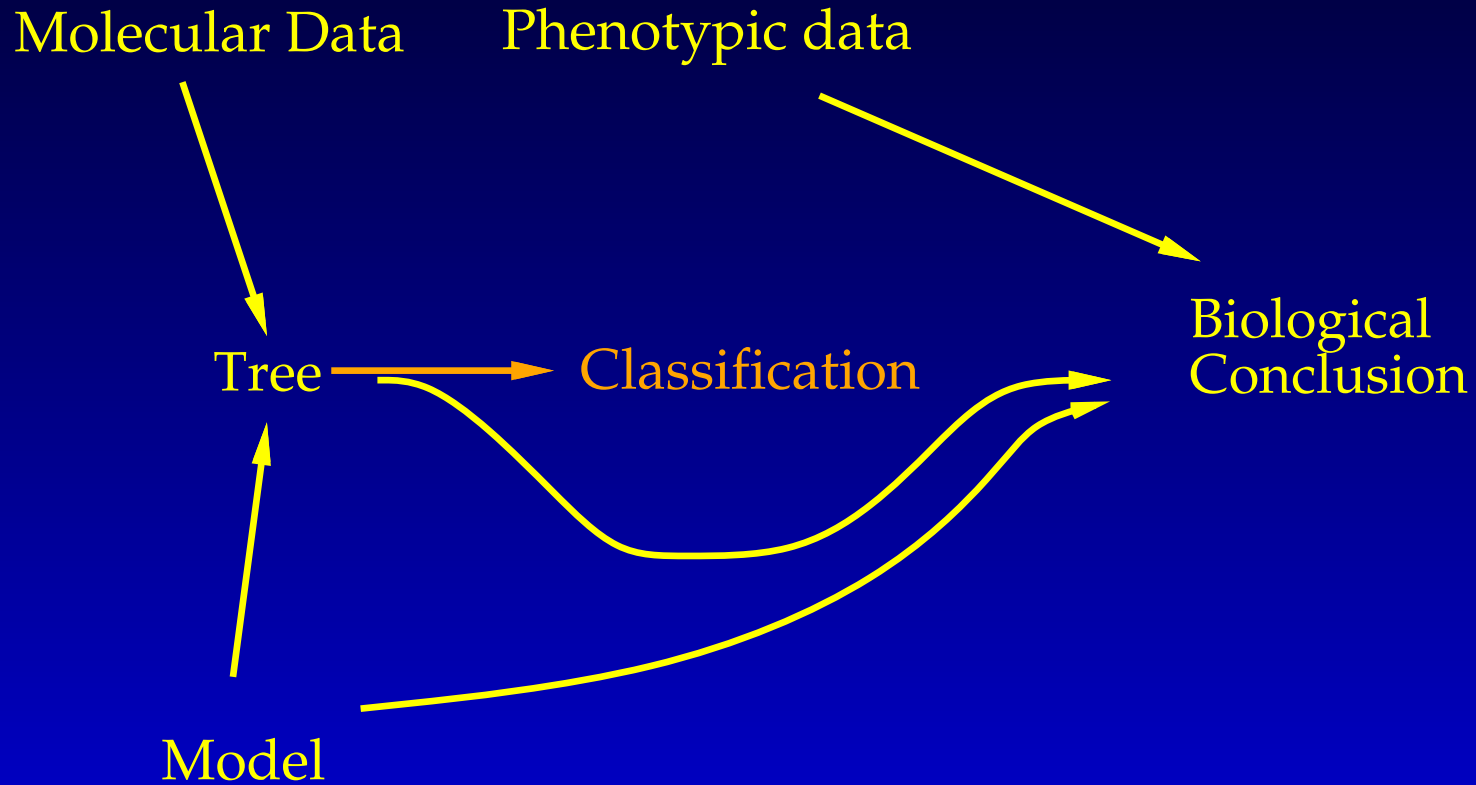


Model

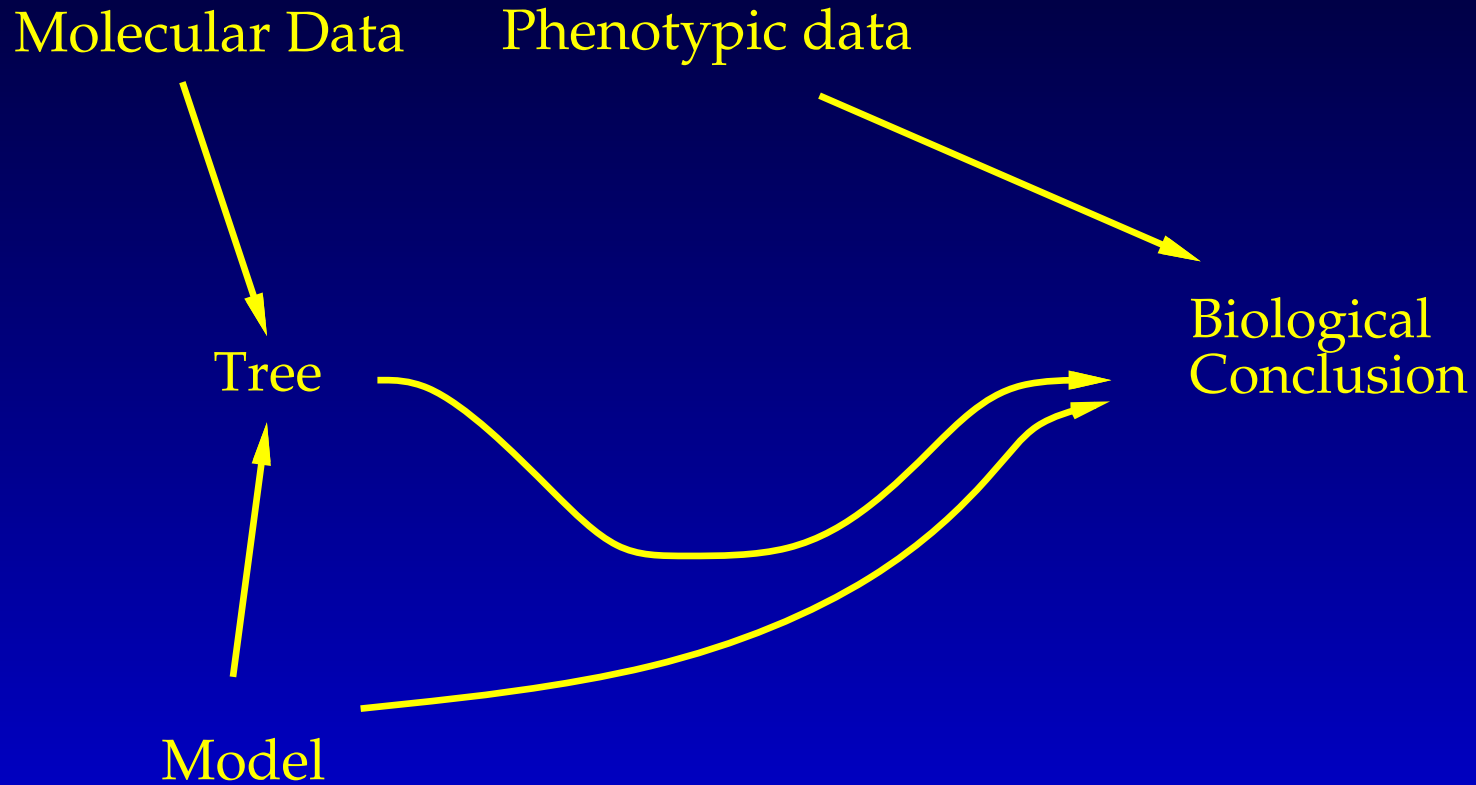
Now: the importance of (beta) classification?



Now: the importance of (beta) classification?



Future? the unimportance of (beta) classification



“But you need names to be able to talk about organisms!”

Yes, but ...

- Why then do we need names that describe nonoverlapping sets?
- Why require monophyly?
- In short, does the need for names specify that we must have a hierarchical classification?

What would Julian Huxley have thought?



Julian Huxley as Fellow of New College, Oxford, 1922

What would Julian Huxley have ordered?



The Sir Julian Huxley, 152-154 Addington Road, Selston, Surrey

References for the threshold method

- Wright, S. 1934. An analysis of variability in number of digits in an inbred strain of guinea pigs. *Genetics* 19: 506-536. [The threshold model for discrete traits]
- Falconer, D. S. 1965. The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Annals of Human Genetics* 29: 51-76. [Threshold model applied to human diseases]
- Lewis, P. O. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology* 50: 913-925. [Uses 0/1 stochastic process to infer morphological phylogenies]
- Pagel, M. 1994. Detecting correlated evolution on phylogenies: A general method for the comparative analysis of discrete characters. *Proceedings of the Royal Society of London Series B Biological Sciences*. 255: 37-45. [0/1 stochastic model for discrete characters]
- Felsenstein, J. 2005. Using the quantitative genetic threshold model for inferences between and within species. *Philosophical Transactions of the Royal Society of London, series B* 360: 1427-1434. [This project in a slightly earlier version]

How it was done

This presentation was prepared using freeware:

- LaTeX (mathematical typesetting and PDF preparation)
- prosper class for projection slides
- Idraw (drawing program to modify plots and draw figures)
- dvips to prepare Postscript file
- ps2pdf to turn this into a PDF
- Adobe Acrobat Reader (to display the PDF in full-screen mode)
- Linux (operating system)