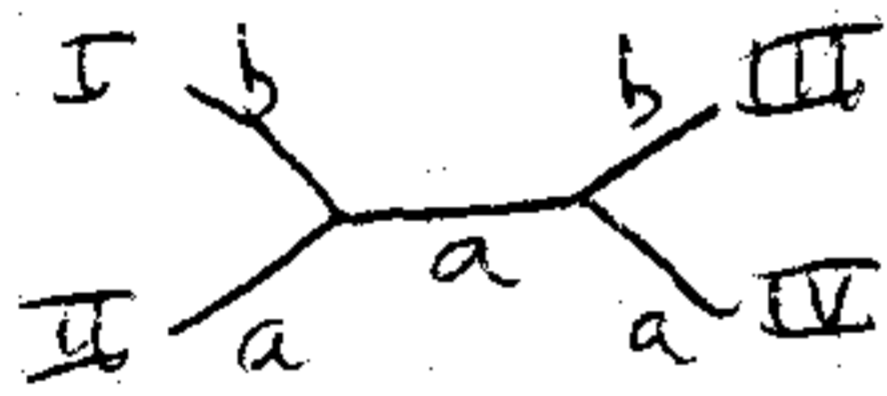


JOHN HUELSENBECK

Criteria

- consistent = converges to correct tree w/ sufficient data
- efficient = quickly converges on correct tree
- robust = correctly estimated phylogeny even if assumptions are violated
- speed
- falsifiability =
- versatility =

Problems w/ simulation
- validation?



Parameters

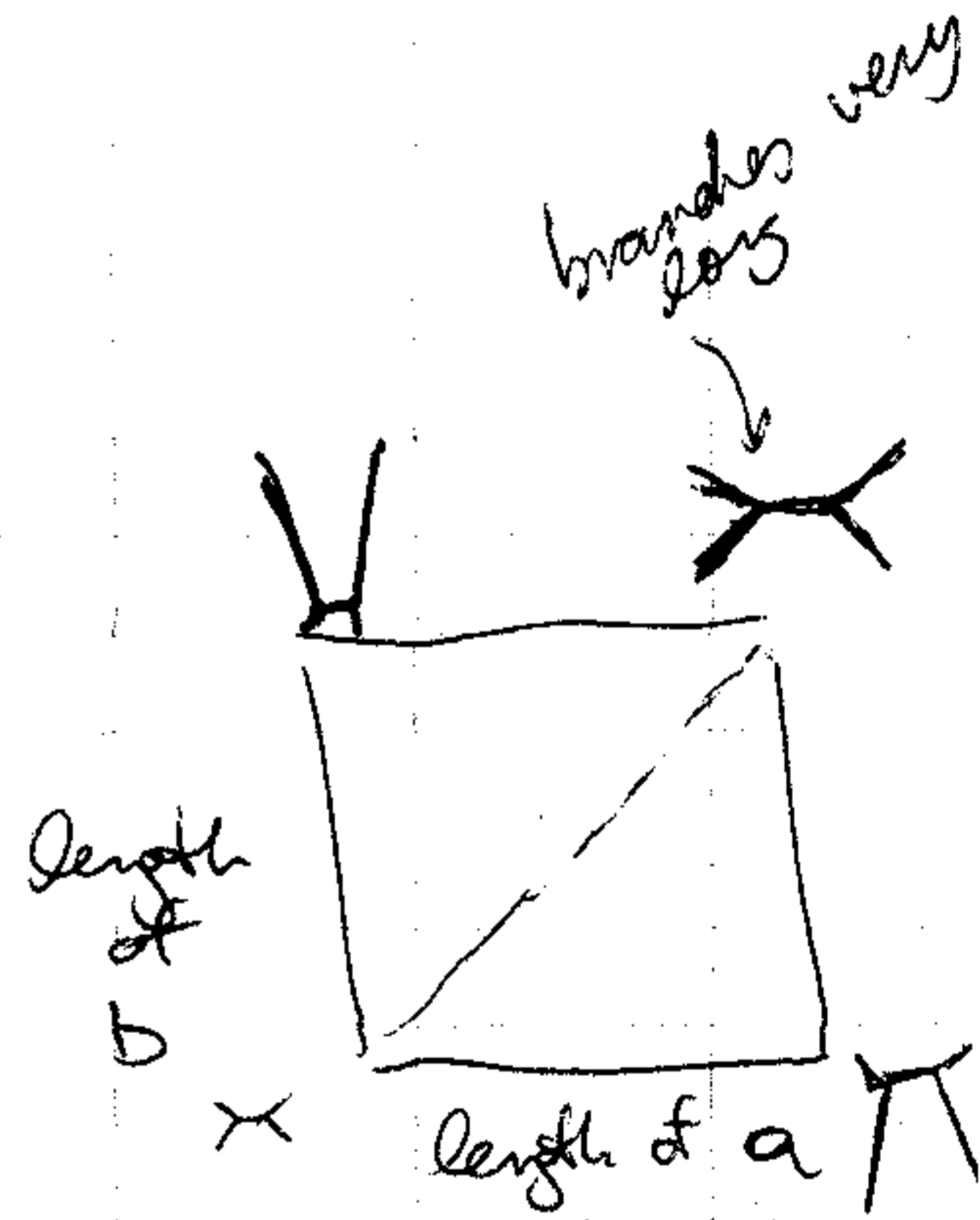
⊙ branch length variation

Results

Felsenstein zone = low performance



does OK here



Parsimony

- Felsenstein zone is due to convergence at tips
- because using 4 taxa parsimony can only infer a tree if tips are identical to each other

Parametric Bootstrap

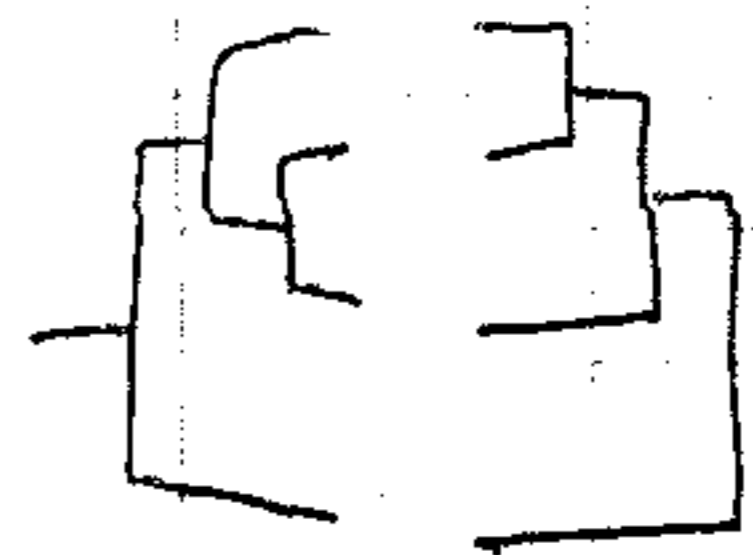
- ① take tree
- ② infer subst. pattern
- ③ re-evolve characters
- ④ ~~take~~ which tree do you get?

How likely are long branch problems?

- weighting improves performance
- model most important w/ rate variation
- low variance good
- ML & addition distance methods best
- UPGMA limited

How test?

- ① take well supported trees
- ② take genes from these species
- ③ use boots & multiple methods



Weighting?

Making a step matrix

- count S's

Diff. species