

dmeav02 Usage

Francisco Richter M.

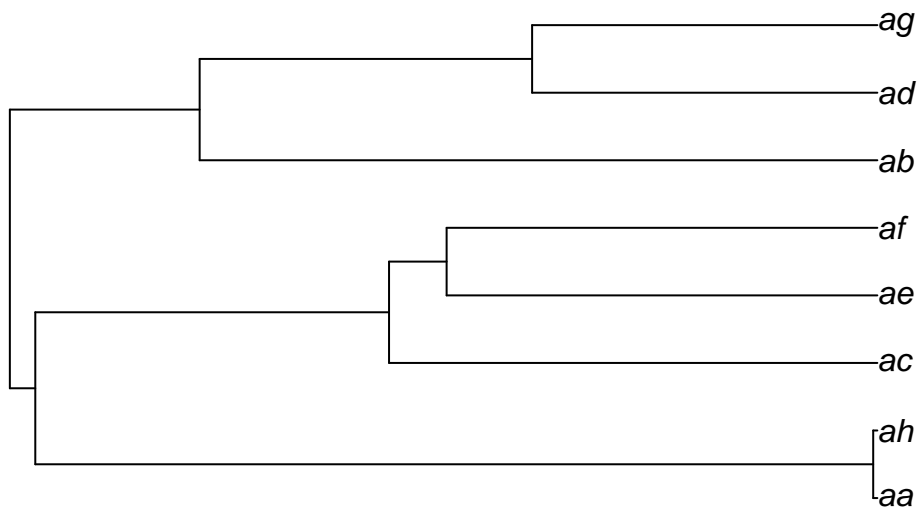
June 1, 2017

tree simulation and consistency

functions

1. `sim.tree` to simulate trees
2. `phylo2vectors` to convert a phylo into the set of vectors characterizing the tree

```
s = sim.tree(ct=6)
plot(s$phylo)
```



```
s$tree
```

```
## $wt
## [1] 3.14447149 0.08380073 0.54094806 0.62341551 0.18957634 0.28140727
## [7] 1.12273065 0.01364995
##
## $E
## [1] 1 1 1 1 1 1 1
##
## $n
## [1] 1 2 3 4 5 6 7 8
##
## $S
## [1] "aa" "aa" "ab" "ac" "ae" "ad" "aa"
##
## $br
## [1] 3.144471 3.228272 3.769220 4.392636 4.582212 4.863619 5.986350 6.000000
```

```
phylo2vectors(s$phylo)
```

```
## $wt
## [1] 3.14447149 0.08380073 0.54094806 0.62341551 0.18957634 0.28140727
## [7] 1.12273065 0.01364995
```

```
##
## $E
## [1] 1 1 1 1 1 1 1
##
## $n
## [1] 1 2 3 4 5 6 7 8
all.equal(s$tree$wt,phylo2vectors(s$phylo)$wt)
```

```
## [1] TRUE
all.equal(s$tree$E,phylo2vectors(s$phylo)$E)
```

```
## [1] TRUE
all.equal(s$tree$n,phylo2vectors(s$phylo)$n)
```

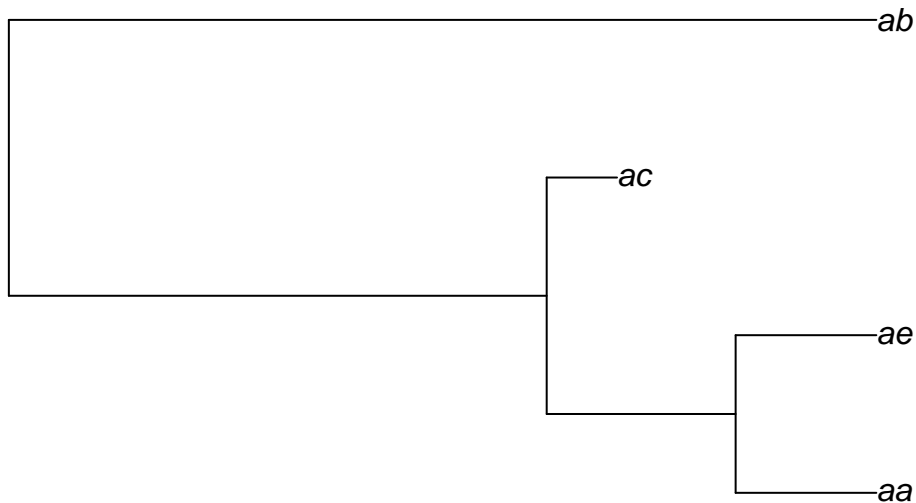
```
## [1] TRUE
```

to do:

1. phylo2vectors does not give the vector S corresponding to topology. add it
2. remove the root on vector outputs, keep consistency
3. arrange the vectors2phylo function and make a lot simpler the `sim.tree` using it
4. add the log in an log output file

Tree manipulation

```
s = sim.tree(ct=2,seed=5)
plot(s$phylo)
```



```
s2=phylo2vectors(s$phylo)
s2
```

```
## $wt
## [1] 0.1341027 1.1568185 0.1512854 0.2551456 0.3026478
##
## $E
## [1] 1 1 0 1
##
```

```
## $n
## [1] 1 2 3 2 3
```

```
s$tree
```

```
## $wt
## [1] 0.1341027 1.1568185 0.1512854 0.2551456 0.3026478
##
## $E
## [1] 1 1 0 1
##
## $n
## [1] 1 2 3 2 3
##
## $S
## [1] "aa" "aa" "ac" "aa"
##
## $br
## [1] 0.1341027 1.2909212 1.4422066 1.6973522 2.0000000
```

```
up = update.tree(s2,0.5,1.5)
```

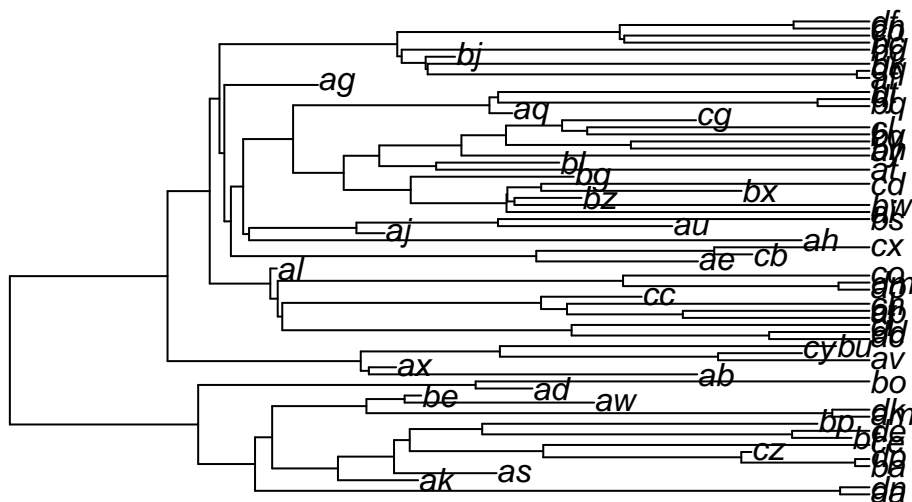
```
up
```

```
## $wt
## [1] 0.1341027 0.3658973 0.7909212 0.1512854 0.0577934 0.1973522 0.3026478
##
## $E
## [1] 1 1 1 0 0 1
##
## $n
## [1] 1 2 3 4 3 2 3
```

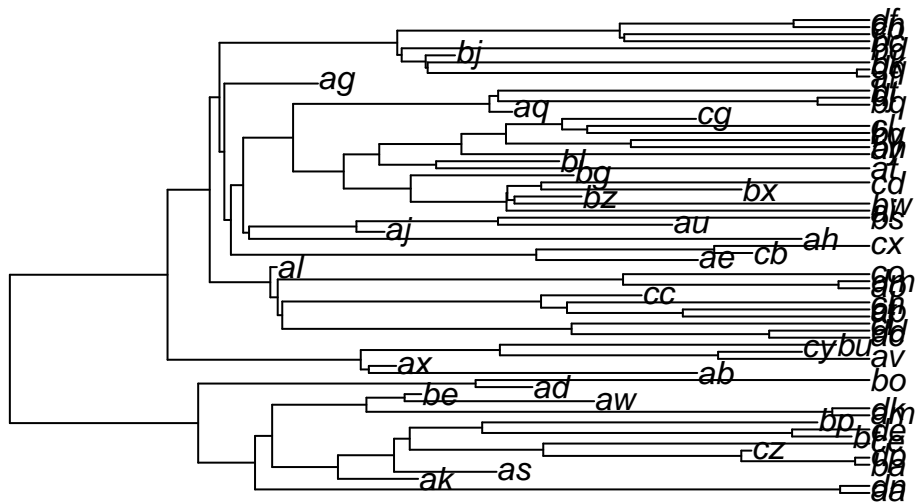
1. Doing update tree I realized that I should be more carefully in analysis when we lost crown time.
2. phylo2vectors really needs to take into account topology

more about consistency

```
s <- sim.tree()
plot(s$phylo)
```



```
plot(vectors2phylo(s$tree))
```

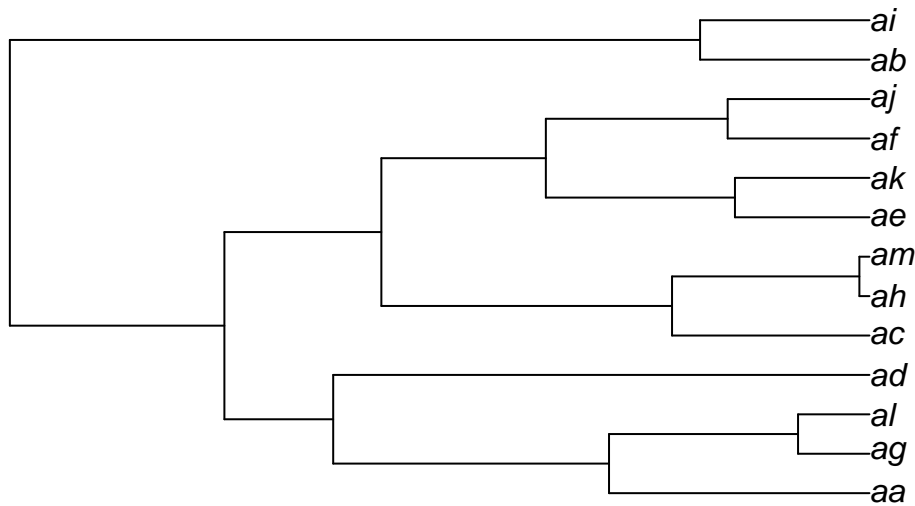


```
all.equal(s$phylo,vectors2phylo(s$tree))
```

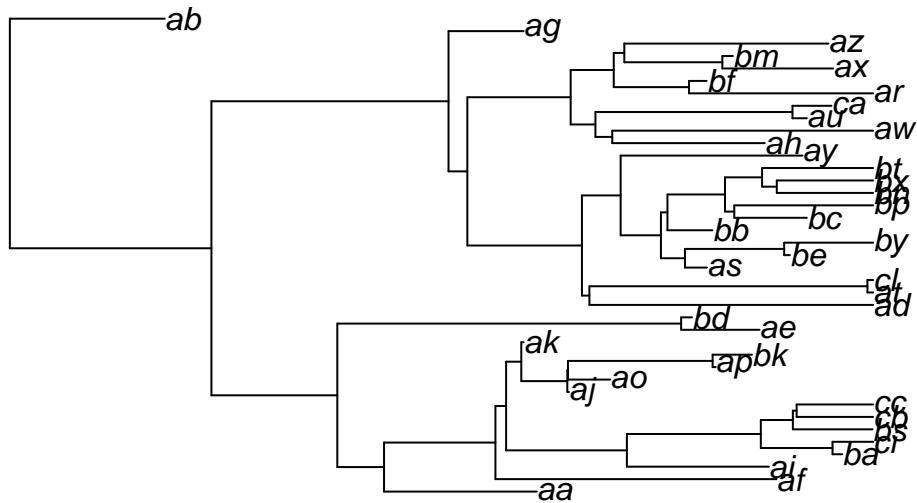
```
## [1] TRUE
```

The reconstruction algorithm

```
s = sim.tree(ct=6)
plot(s$phylo.extant)
```



```
rec = rec.tree(tree=s$tree.extant,pars=c(0.8,0.1,40))
plot(vectors2phylo(rec))
```



```
s$tree.extant
```

```
## $wt
## [1] 3.09113043 0.72604639 0.36817361 0.16296890 0.55666225 0.21378666
## [7] 0.21317853 0.09490394 0.09335269 0.02462513 0.21356805 0.20749833
## [13] 0.03410509
##
```

```
## $E
## [1] 1 1 1 1 1 1 1 1 1 1 1 1 1
##
```

```
## $n
## [1] 1 2 3 4 5 6 7 8 9 10 11 12 13
```

```
rec
```

```
## $wt
## [1] 0.8657115491 0.9210890817 0.2777939960 0.7486506488 0.2778851552
## [6] 0.3834168296 0.1120796602 0.1669579777 0.0635919236 0.0901771714
## [11] 0.0139233123 0.0004398246 0.0783482260 0.1810983355 0.0041867355
## [16] 0.0053875897 0.0101847897 0.0672105589 0.0443074560 0.0358785021
## [21] 0.0880994090 0.0119551983 0.0089233092 0.0410880943 0.0216384357
## [26] 0.0154769913 0.2022449266 0.0389254772 0.0813799539 0.0226468212
## [31] 0.0242836334 0.0172536121 0.0878701462 0.0018450112 0.0320001971
## [36] 0.0014500506 0.0179018053 0.0399223447 0.0155434887 0.0471731545
## [41] 0.0081300776 0.1051804752 0.0445463049 0.0081485161 0.0076354493
## [46] 0.0158568039 0.0243782991 0.0458452905 0.0011880946 0.0440415326
## [51] 0.0338081485 0.0155030114 0.0022699734 0.0223551605 0.0352826092
## [56] 0.0251685027 0.0004655672 0.1287736902 0.0175990836 0.0062786016
## [61] 0.0038777863 0.0540461950 0.1495743488 0.0305880531 0.0035170407
##
## $E
## [1] 1 0 1 1 1 1 1 1 1 1 0 0 0 1 1 0 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 0 0 0
## [36] 1 0 1 1 0 1 0 0 1 1 0 0 0 1 1 0 1 1 1 0 0 0 0 0 1 0 0 1 0
##
## $n
## [1] 1 2 1 2 3 4 5 6 7 8 9 8 7 6 7 8 7 8 9 10 11 10 11
## [24] 12 13 14 15 16 17 18 19 20 19 18 17 16 17 16 17 18 17 18 17 16 17 18
## [47] 17 16 15 16 17 16 17 18 19 18 17 16 15 14 15 14 13 14 13
```

let's observe one estimation

```
s = sim.tree(seed = 1)
tree = s$tree.extant
st = sim.srt(tree,pars=c(0.8,0.1,40),n_trees=100)
mle.st(st)
```

```
## [1] 0.9805493 0.1234330 42.8584263
```

How good is the last iteration of the MCEM algorithm?

```
library(dmea)
n_sim = 537
n_trees = 10
MP = matrix(nrow=n_sim,ncol=3)
RP = matrix(nrow=n_sim,ncol=3)
p = proc.time()
for(i in 1:n_sim){
  est = sim.est(n_trees=n_trees,pars=c(0.8,0.1,40),seed=i)
  RP[i,] = est$real
  MP[i,] = est$est
}
print(proc.time()-p)
par_est_vis(P=MP,par=1,PR=RP)
par_est_vis(P=MP,par=2,PR=RP)
par_est_vis(P=MP,par=3,PR=RP)
```