

# Package ‘phylopath’

October 5, 2016

**Type** Package

**Title** Perform Phylogenetic Path Analysis

**Version** 0.2.0

**Author** Wouter van der Bijl

**Maintainer** Wouter van der Bijl <wouter.van.der.bijl@zoologi.su.se>

**Description** A comprehensive and easy to use R implementation of confirmatory phylogenetic path analysis as described by Von Hardenberg and Gonzalez-Voyer (2012) <doi:10.1111/j.1558-5646.2012.01790.x>.

**URL** <https://github.com/Ax3man/phylopath>

**BugReports** <https://github.com/Ax3man/phylopath/issues>

**License** GPL-3

**LazyData** TRUE

**Depends** R (>= 2.10)

**Imports** ape, DiagrammeR, dplyr, ggm, ggplot2, igraph, MuMIn, nlme, purrr, tibble, tidyr

**RoxygenNote** 5.0.1

**Suggests** knitr, rmarkdown

**VignetteBuilder** knitr

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2016-10-05 09:29:55

## R topics documented:

average	2
average_DAGs	3
best	4
coef_plot	5
DAG	5

est_DAG . . . . .	6
phylo_path . . . . .	7
rhino . . . . .	8
rhino_tree . . . . .	8

<b>Index</b>	<b>9</b>
--------------	----------

---

average	<i>Extract and average the best supported models for a phylogenetic path analysis.</i>
---------	--

---

## Description

Extract and average the best supported models for a phylogenetic path analysis.

## Usage

```
average(phylopath, cut_off = 2, method = "conditional", ...)
```

## Arguments

phylopath	An object of class phylopath.
cut_off	The CICc cut-off used to select the best models. Use Inf to average over all models. Use the best function to only use the top model.
method	Either "full" or "conditional". The methods differ in how they deal with averaging a path coefficient where the path is absent in some of the models. The full method sets the coefficient (and the variance) for the missing paths to zero, meaning paths that are missing in some models will shrink towards zero. The conditional method only averages over models where the path appears, making it more sensitive to small effects. Following von Hardenberg & Gonzalez-Voyer 2013, conditional averaging is set as the default. Also see <a href="#">model.avg</a> .
...	Additional arguments passed to <a href="#">par.avg</a> . For details on the error calculations, see <a href="#">par.avg</a> .

## Value

An object of class fitted\_DAG.

## Examples

```

candidates <- list(A = DAG(LS ~ BM, NL ~ BM, DD ~ NL + LS),
                  B = DAG(LS ~ BM, NL ~ LS, DD ~ NL),
                  C = DAG(LS ~ BM, NL ~ LS + BM, DD ~ NL))
p <- phylo_path(candidates, rhino, rhino_tree)
summary(p)
# Models A and C have close to equal support, so we may decide to take
# their average.

```

```

avg_model <- average(p)
# Print the average model to see coefficients, se and ci:
avg_model
# Plot to show the weighted graph:
plot(avg_model)
# Note that coefficients that only occur in one of the models become much
# smaller when we use full averaging:
coef_plot(avg_model)
coef_plot(average(p, method = 'full'))

```

---

average\_DAGs

*Perform model averaging on a list of DAGs.*


---

## Description

Perform model averaging on a list of DAGs.

## Usage

```

average_DAGs(fitted_DAGs, weights = rep(1, length(coef)),
  method = "conditional", ...)

```

## Arguments

fitted_DAGs	A list of fitted_DAG objects containing coefficients and standard errors, usually obtained by using <code>est_DAG</code> on several DAGs.
weights	A vector of associated model weights.
method	Either "full" or "conditional". The methods differ in how they deal with averaging a path coefficient where the path is absent in some of the models. The full method sets the coefficient (and the variance) for the missing paths to zero, meaning paths that are missing in some models will shrink towards zero. The conditional method only averages over models where the path appears, making it more sensitive to small effects. Following von Hardenberg & Gonzalez-Voyer 2013, conditional averaging is set as the default. Also see <a href="#">model.avg</a> .
...	Additional arguments passed to <a href="#">par.avg</a> . For details on the error calculations, see <a href="#">par.avg</a> .

## Value

An object of class `fitted_DAG`, including standard errors and confidence intervals.

**Examples**

```
# Normally, I would advocate the use of the phylo_path and average
# functions, but this code shows how to average any set of models. Note
# that not many checks are implemented, so you may want to be careful and
# make sure the DAGs make sense and contain the same variables!
candidates <- list(A = DAG(LS ~ BM, NL ~ BM, DD ~ NL),
                  B = DAG(LS ~ BM, NL ~ LS, DD ~ NL))
fit_cand <- lapply(candidates, est_DAG, rhino, ape::corPagel, rhino_tree)
ave_cand <- average_DAGs(fit_cand)
coef_plot(ave_cand)
```

---

 best

*Extract and estimate the best supported model for a phylogenetic path analysis.*

---

**Description**

Extract and estimate the best supported model for a phylogenetic path analysis.

**Usage**

```
best(phylopath)
```

**Arguments**

phylopath      An object of class phylopath.

**Value**

An object of class fitted\_DAG.

**Examples**

```
candidates <- list(A = DAG(LS ~ BM, NL ~ BM, DD ~ NL),
                  B = DAG(LS ~ BM, NL ~ LS, DD ~ NL))
p <- phylo_path(candidates, rhino, rhino_tree)
best_model <- best(p)
# Print the best model to see coefficients, se and ci:
best_model
# Plot to show the weighted graph:
plot(best_model)
```

---

coef_plot	<i>Plot path coefficients and their confidence intervals.</i>
-----------	---

---

**Description**

Plot path coefficients and their confidence intervals.

**Usage**

```
coef_plot(fitted_DAG, reverse_order = FALSE)
```

**Arguments**

`fitted_DAG` A fitted DAG, usually obtained by `best`, `average` or `est_DAG`.

`reverse_order` If TRUE, the paths are plotted in reverse order. Particularly useful in combination with `ggplot2::coord_flip` to create horizontal versions of the plot.

**Value**

A ggplot object.

**Examples**

```
d <- DAG(LS ~ BM, NL ~ BM, DD ~ NL + LS)
plot(d)
d_fitted <- est_DAG(d, rhino, ape::corBrownian, rhino_tree)
plot(d_fitted)
coef_plot(d_fitted)
# to create a horizontal version, use this:
coef_plot(d_fitted, reverse_order = TRUE) + ggplot2::coord_flip()
```

---

DAG	<i>Directed acyclic graphs (DAGs)</i>
-----	---------------------------------------

---

**Description**

This function is a simple wrapper around the function from the `ggm` package. The only differences are that the `order` argument defaults to TRUE and that it adds a DAG class for easy plotting.

**Usage**

```
DAG(..., order = TRUE)
```

**Arguments**

... a sequence of model formulae

order logical, defaulting to FALSE. If TRUE the nodes of the DAG are permuted according to the topological order. If FALSE the nodes are in the order they first appear in the model formulae (from left to right).

**Value**

An object of classes matrix and DAG

**Examples**

```
# Use formula notation to create DAGs:
plot(DAG(A~B, B~C))
# Use + to easily add multiple parents to a node:
plot(DAG(A~B+C))
# Add a node as it's own parent to create an isolate:
plot(DAG(A~B+C, D~D))
```

---

est\_DAG

*Add standardized path coefficients to a DAG.*

---

**Description**

Add standardized path coefficients to a DAG.

**Usage**

```
est_DAG(DAG, data, cor_fun, tree)
```

**Arguments**

DAG A directed acyclic graph, typically created with DAG.

data A data.frame with data.

cor\_fun A function that creates a corStruct object, typically one of the cor function from the ape, such as corBrownian, corPagel etc.

tree A phylogenetic tree of class pylo.

**Value**

An object of class fitted\_DAG.

**Examples**

```
d <- DAG(LS ~ BM, NL ~ BM, DD ~ NL + LS)
plot(d)
d_fitted <- est_DAG(d, rhino, ape::corBrownian, rhino_tree)
plot(d_fitted)
```

---

 phylo\_path

*Compare causal models in a phylogenetic context.*


---

### Description

Compare causal models in a phylogenetic context.

### Usage

```
phylo_path(models, data, tree, order = NULL, cor_fun = ape::corPagel)
```

### Arguments

models	A list of directed acyclic graphs. These are matrices, typically created with DAG.
data	A data.frame with data.
tree	A phylogenetic tree of class <code>pylo</code> .
order	Causal order of the included variable, given as a character vector. This is used to determine which variable should be the dependent in the dsep regression equations. If left unspecified, the order will be automatically determined. If the combination of all included models is itself a DAG, then the ordering of that full model is used. Otherwise, the most common ordering between each pair of variables is used to create a general ordering.
cor_fun	A function that creates a <code>corStruct</code> object, typically one of the <code>cor</code> function from the <code>ape</code> , such as <code>corBrownian</code> , <code>corPagel</code> etc.

### Value

A table with relevant statistics for each model, including the C statistic, the associated p-values, information criterions, and model weights.

### Examples

```
#see vignette('intro_to_phylopath') for more details
candidates <- list(A = DAG(LS ~ BM, NL ~ BM, DD ~ NL),
                  B = DAG(LS ~ BM, NL ~ LS, DD ~ NL))
p <- phylo_path(candidates, rhino, rhino_tree)

# Printing p gives some general information:
p
# And the summary gives statistics to compare the models:
summary(p)
```

---

rhino

*Rhinogrades traits*

---

**Description**

A simulated dataset, as used by Gonzalez-Voyer and Von Hardenberg as an example, containing variables on body mass (BM), litter size (LS), nose length (NL), dispersal distance (DD) and range size (RS).

**Usage**

rhino

**Format**

An object of class `data.frame` with 100 rows and 6 columns.

**Source**

> Gonzalez-Voyer A & von Hardenberg A. 2014. An Introduction to Phylogenetic Path Analysis. Chapter 8. In: Garamszegi LZ (ed.), Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology. pp. 201-229. Springer-Verlag Berlin Heidelberg. <doi:10.1111/j.1558-5646.2012.01790.x>

---

rhino\_tree

*Rhinogrades phylogeny*

---

**Description**

A phylogenetic tree for the 100 species of the rhino dataset.

**Usage**

rhino\_tree

**Format**

An object of class `phylo` of length 4.

**Source**

> Gonzalez-Voyer A & von Hardenberg A. 2014. An Introduction to Phylogenetic Path Analysis. Chapter 8. In: Garamszegi LZ (ed.), Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology. pp. 201-229. Springer-Verlag Berlin Heidelberg. <doi:10.1111/j.1558-5646.2012.01790.x>



# Index

## \*Topic **datasets**

rhino, 8

rhino\_tree, 8

average, 2

average\_DAGs, 3

best, 4

coef\_plot, 5

DAG, 5

est\_DAG, 6

model.avg, 2, 3

par.avg, 2, 3

phylo\_path, 7

rhino, 8

rhino\_tree, 8