

# Tutorial for DAGitty

Brice Ozenne

This document describes DAGitty, a software that can be used to:

- draw DAGs
- indicates the observed and unobserved variables
- decide which on which set(s) of variables one should adjust on.

It is a free software that can be downloaded at <http://www.dagitty.net/>. Note that there is also a corresponding R package called `dagitty` available on CRAN. For this tutorial we will use the web interface that can be access using <http://www.dagitty.net/dags.html>. When opening the page you should see:

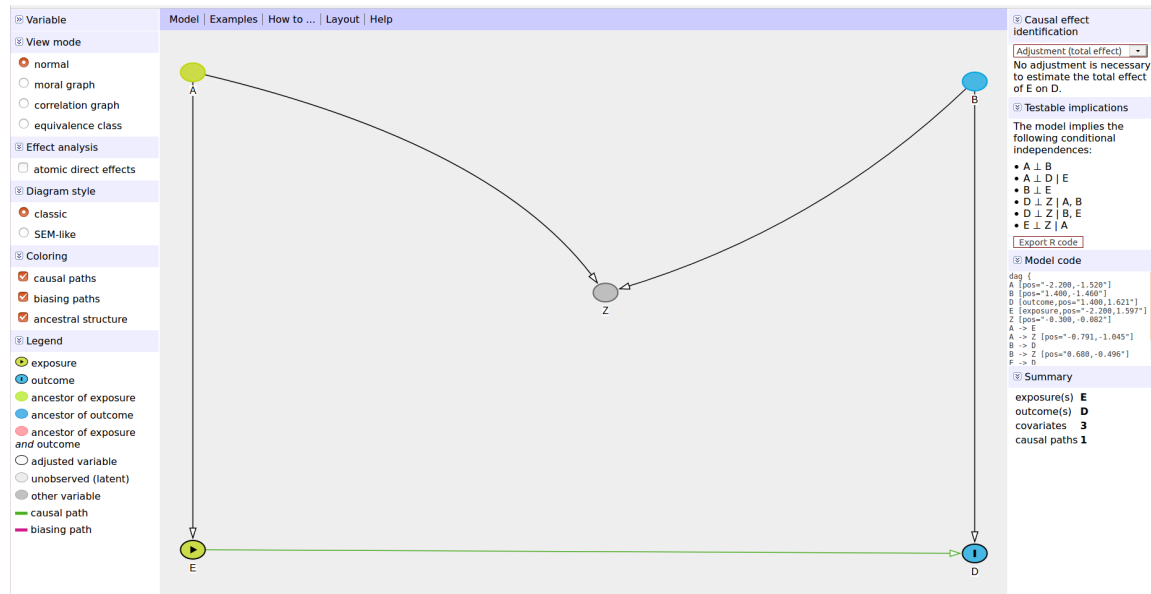




Figure 1: Home page of the web interface of DAGitty

# 1 User interface

We will only present some of the functionalities of DAGitty. In the web interface, we will:

- uncheck the box "ancestral structure" in the sub-menu "Coloring" in the left column.
- click on the double arrow  to reduce the "View mode", "Effect analysis", "Diagram style", "Coloring" in the left column, and "Testable implications" in the right column.
- click on the double arrow  to expand the "Variable" menu (left column)

You should now see something similar as [Figure 2](#).

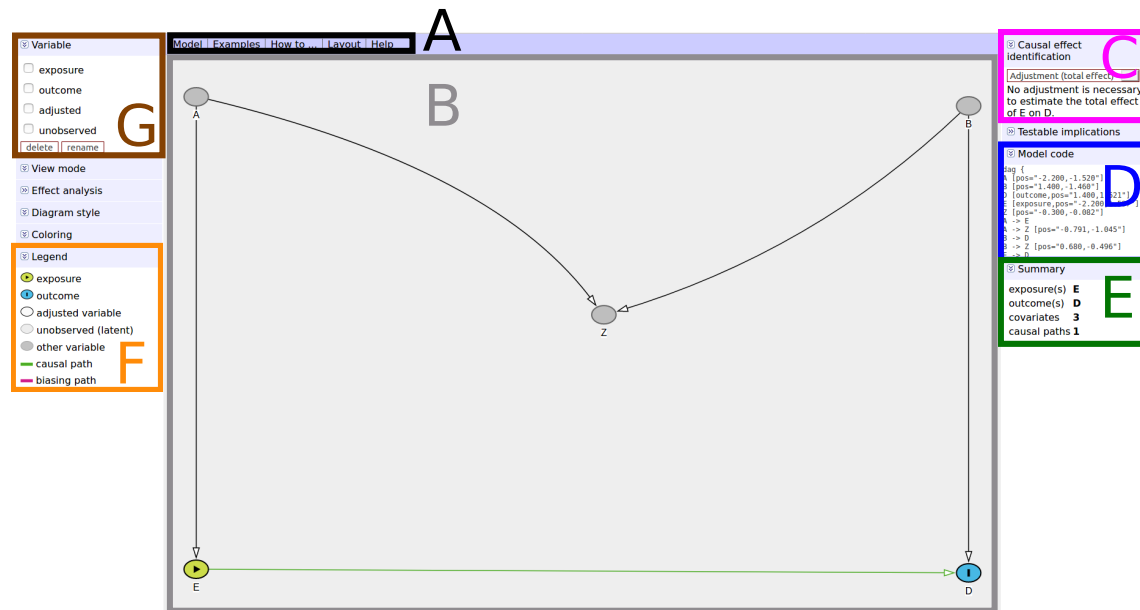


Figure 2: Key elements of the web interface of DAGitty

We can divide the display into 7 blocks:

- A: this is the menu where you can get some help about how to do things (see "How to") or export the DAG (see "Model").
- B: contains the DAG. You can interact with it by:
  - clicking in any empty space enables you to **add a node**.
  - clicking on a node and drag it. This enables you to change the display of the dag without affecting its properties.
  - clicking on a node and use the F-menu/specific key binding to **change the type of the node** (e.g. define it as an exposure or unobserved variable) or **remove the node**.

- clicking on a node and then another node to **remove/add an arrow**.
- **C ("Causal effect identification")**: displays whether by adjusting on a specific set of variables, it is possible to estimate the causal effect of the exposure(s) on the outcome. Note that there may be several possible sets to "adjust on" or none (meaning that one cannot estimate the total causal effect). DAGitty will display the "minimal" set of variables: in the example it says that no adjustment is necessary even though adjusting on A and/or B would also be valid.
- **D ("Model code")**: contains the code used to generate the graph. It is a way to save the graph: by copying the code somewhere on your computer you can, next time you launch DAGitty, paste it in the "Model code" section and that will generate your previous graph.
- **E ("Summary")**: summary of the key elements of the DAGs. It for instance lists the number of causal paths from the exposure to the outcome (here only one causal path, the arrow between E and D).
- **F ("Legend")**: explains how each type of variable is displayed, e.g. the outcome is displayed with a blue circle and a vertical bar at the center and unobserved variables are displayed in dark gray (here A, Z, and B). Causal paths will be displayed in green while non-causal paths will be displayed in pink.
- **G ("Variable")**: enables to update the type of a node (i.e. a variable) or remove it. To use it you need first to select the node.

## 2 Example

To create a new DAG you can:

1. click on the "Model" menu and on the "New model" in the list.
2. click anywhere in the empty space (panel B). That will begin the creation of a node (i.e. a variable).
3. Type the name of the variable in the box (e.g. remove `v1` and type `treatment`) and click ok.
4. The node has been created!
5. Perform the same operation to create a second node, e.g. called cancer.
6. Select one node (e.g. `treatment`) and the select another node (e.g. `cancer`).
7. This created an arrow from the first to the second node.

8. Select a node (e.g. treatment) and then type "e" or go to the "Variable" (panel G) and select exposure.
9. The node is now considered as an exposure.

You can continue:

- identifying the other node as an outcome
- add other variables and arrows relating variables
- indicate the unobserved variables
- indicate the variables you adjust on by design (e.g. in a case-control study, sampling cases among hospital in-patient population condition on the hospitalization node, with an arrow from the outcome to the hospitalization node).
- "Causal effect identification" (panel C) will indicate whether it is possible to identify the total causal effect of the exposure on the outcome. If yes it will indicate on which set of variables should we adjust on in the analysis.

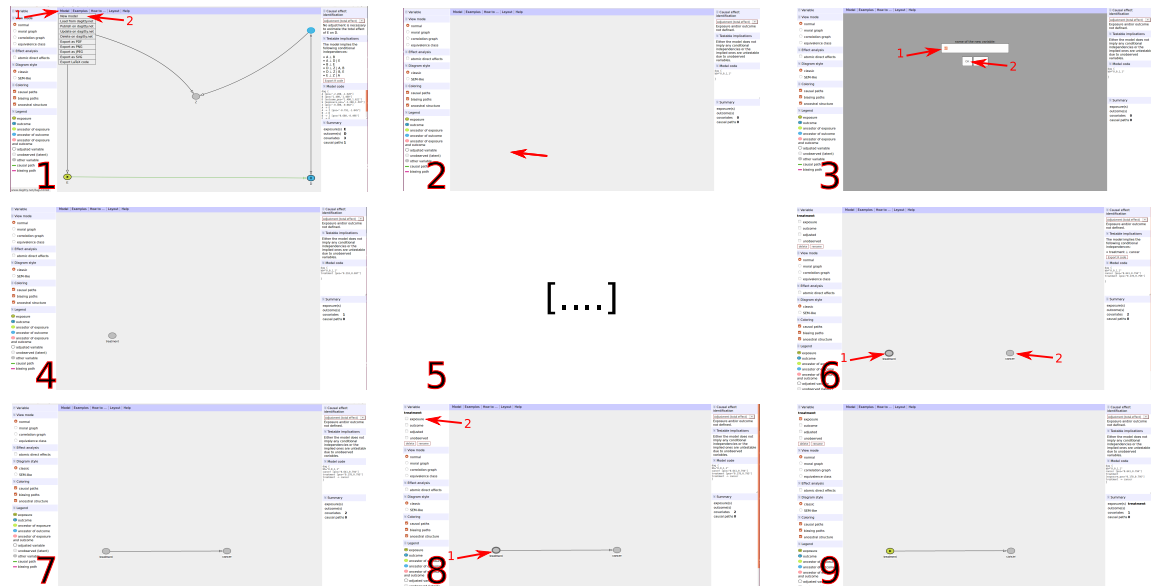


Figure 3: Construction of a DAG using DAGitty

### 3 Key binding

- n: add a new node (i.e. a new variable)

To use the following key binding, You need first to have clicked on a node (i.e. an existing variable).

- e: turn a variable into an exposure.
- o: turn a variable into an outcome.
- a: condition on this variable. It becomes an "adjusted" variable.
- u: turn a variable into an unobserved variable (also called latent variable).
- r: rename a node (i.e. a new variable)
- d: delete a variable.