Ralph's DAG

Ralph knows DAG (directed acyclic graph) structures are convenient and instructive frames for structural causal models (SCM). Directions in the graph depict causal relations amongst observables (as well as with unobservables). The ideas are summarized in the following definition of d-separation and theorem connecting d-separation with conditional independence.

Definition 1. *[d-separation] a path p is d-separated (blocked) by a set of nodes Z (including the null set $\varnothing$) if and only if*
1. p contains a chain $i \to m \to j$ or fork $i \leftarrow m \to j$ such that the middle node m is in Z, or
2. p contains an inverted fork (collider) $i \to m \leftarrow j$ such that the middle node m is not in Z and no descendant of m is in Z.

A set Z d-separates X and Y if and only if Z blocks every path from X to Y.

Theorem 2. *[d-separation and conditional independence] If sets X and Y are d-separated by Z in a DAG G, then X is independent of Y conditional on Z in every distribution consistent with G. Conversely, if X and Y are not d-separated by Z in a DAG G, then X and Y are dependent conditional on Z in at least one distribution consistent with G.*

The converse part is actually much stronger. If X and Y are not blocked then they are dependent in almost all distributions consistent with G. Independence of unblocked paths requires precise parameter tuning that is unlikely.

Ralph constructs a DAG utilizing his understanding of the setting along with conditional independence tests of (in principle) all subsets of variables. Ralph implements conditional independence tests using partial (linear) regression coefficients (but also recognizes, for instance, nonparametric regression could be utilized). Ralph acknowledges all measured variables may not be employed in his empirical design but may nonetheless prove valuable in diagnosing the DAG.

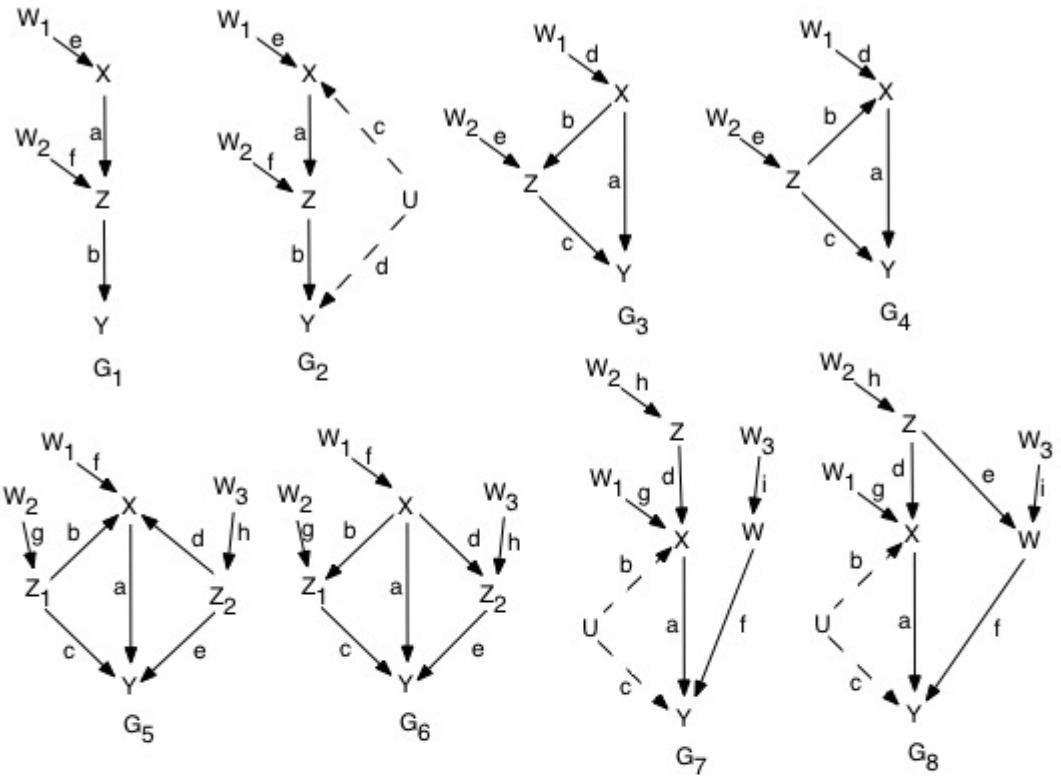
Ralph is attempting to distinguish between pairs of DAGs. That is, G_1 vs. G_2, G_3 vs. G_4, G_5 vs. G_6, and G_7 vs. G_8.

Note: U is unobservable.

Suggested:

1. For each pair, indicate what conditional independence tests are diagnostic.

2. For each DAG, indicate which (observable) variables are to be included in empirical designs exploring the causal effect of X on Y.


W1
e
X
a
W2
f
Z
b
Y
G1
W1
e
X
a
c
W2
f
Z
U
b
d
Y
G2
W1
d
X
b
W2
e
Z
a
c
Y
G3
W1
d
X
b
W2
e
Z
a
c
Y
G4
W1
f
W2
X
W3
g
b
d
h
Z1
a
Z2
c
e
Y
G5
W1
f
W2
X
W3
g
b
d
h
Z1
a
Z2
c
e
Y
G6
W2
h
Z
W3
W1
g
d
i
X
W
b
U
a
f
c
Y
G7
W2
h
Z
W3
W1
g
d
e
i
X
W
b
U
a
f
c
Y
G8