Learning Causal Networks in Python

James Callan

1 Literature Review

1.1 Causal Networks

Causal Networks are graphical models which represent a set of variables, their conditional dependencies, and the causal relationships between them with a Directed Acyclic Graph (DAG). The node of the DAG represent the variables and the edges represent causality, the direction of the edge represents the direction of causality with parent nodes causing child nodes. Causal networks also represent conditional independence with d-separation, variables X and Y are conditionally independent on set Z if nodes X and Y are d-separated by nodes in set Z.

1.2 D Separation

Two nodes are considered unconditionally d-connected if there exists a path between them which is unblocked. They are unconditionally d-separated if the are not unconditionally d-connected. A path is any sequence of adjacent edges regardless of their directionality and unblocked refers to a path that does not traverse a "collider".

A collider is node in a path which is both entered and left on edges which are directed toward the node. If there exists no path between nodes X and Y which does not traverse a collider, nodes X and Y are unconditionally d-separated or d-separated conditioned on the empty set. Two unconditionally d-separated nodes in a causal network are considered to be independent.

Two nodes X and Y are d-separated conditioned on set Z if there exists no path in which all colliders are in set Z or descendants of members of Z. If nodes X and Y are d-separated by conditioning set Z, X is conditionally independent of Y given conditioning set Z.

1.3 Learning Causal Networks

To learn a causal network from data both the causal relationships between variables and conditional independences must be learned. There have been a number of algorithms developed to complete this task. I will be implementing three of them: The PC algorithm, the Fast Causal Inference (FCI) algorithm, and the Really Fast Causal Inference (RFCI) Algorithm. Each of these algorithms begins by computing the skeleton of the DAG.

The skeleton of the DAG contains the edges of the graph however does not specify their directions. To compute