582 2016-02-18

Note Title 2016-02-18

Filling out some conceptual steps.

We define Bayesian networks as done by Neapolitan [1990] (reference on website)

Neapolitan, Richard E. <u>Probabilistic Reasoning in Expert Systems: Theory and Algorithms</u>, Wiley, 1990.

This definition is on the slides. Recall: a Bayesian network is a pair, (G, P), where

G is a DAG with nodes N and directed edges E, and P is a probability distribution on N s.t. every node is conditionally independent of its non-descendants given its parents. [More precise statement on slides]

One can show that d-separation holds for BNs defined as above. Some authors say that "BNs admit d-separation". By this we mean that, if two variables in a BN structure are d-separated by a set of variables, then the same two variables are conditionally independent given the same set of variables. [Can be extended to three sets of variables.]

How does one show that BNs admit d-separation?

In two steps. Theorem 6.1 [Neapolitan]

First, one shows that four axioms (called the graphoid axioms) hold in BNs.

These are the graphoid axioms: (Notation: $I_p(X,Y,Z)$ means that X is indep. of Y given Z.)

1. Symmetry

$$I_p(X,Y,Z) \text{ iff } I_p(Y,X,Z)$$

2. Decomposition

$$I_p(X,Z,Y \cup W) \text{ implies } I_p(X,Z,Y)$$

3. Weak union

$I_p(X, Z, Y \cup W)$ implies $I_p(X, Z \cup Y, W)$

4. Contraction

$I_p(X, Z \cup Y, W)$ and $I_p(X, Z, Y)$ implies $I_p(X, Z, Y \cup W)$.

These axioms were originally stated by Dawid [1979?] and independently rediscovered by Pearl & Paz [1982?]

Second step: One shows that d-seperation follows from the axioms. [Theorem 6.2 [Neapolitan]]

Original proofs are due to Verma & Pearl [1988]; [1990].

Building models; Ch. 3 [JØ7]

Before discussing the good techniques, tricks, etc. in your textbook, one method that always works, but that is not recommended in practice:

the strawman method.

(later)