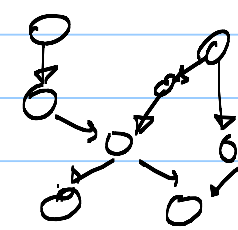
582 2012-01-26

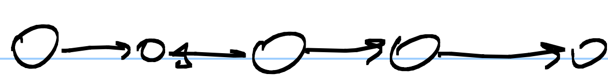
Note Title 2012-01-26

D-seperation (directional separation) (p. 30 [J07])

Defn. Two distinct variables in a causal network (also in a B.N.) are d-seperated [a causal network is just a dag, i.e. a directed acyclic graph (acyclic directed graph, i.e. a directed graph without cycles)]

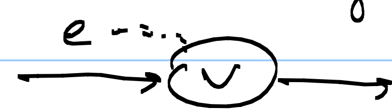


if for all paths [~~better~~ : chains, e.g.



] between A and B, there is an intermediate variable V (distinct from A and B) such that either

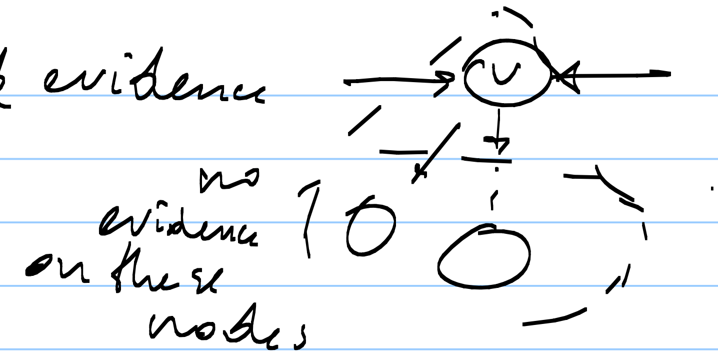
- the connection is serial or diverging and V is instantiated



or

or

- the connection is converging and neither V nor its descendants have received evidence



If A and B are not d-seperated, we call them d-connected.

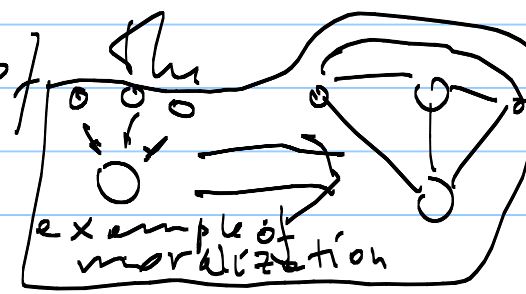
[Note: if A and B are adjacent, they are d-connected; A is d-connected to itself.]

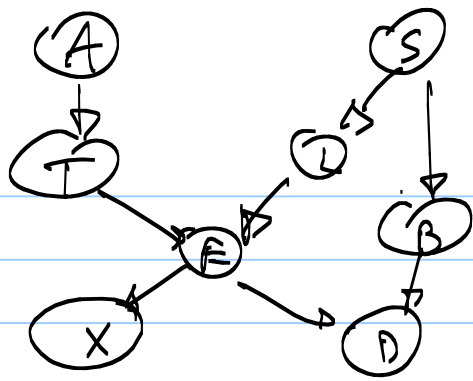
Using the defn. to check d-separation is inefficient, but good enough for small networks.

An efficient algorithm is given on p. 32 [J07]

[Another efficient algorithm is Ross Schacter's Bayes-ball.]

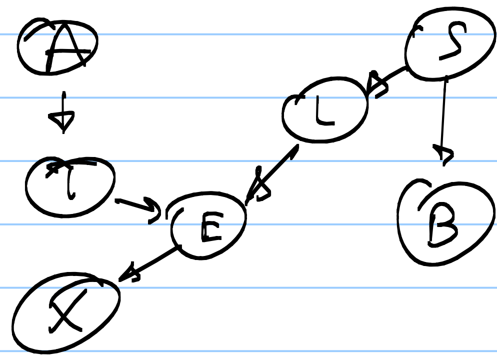
A and B are d-seperated by evidence on node set C iff all (undirected) paths in the (undirected) moral graph of the ancestral graph of A, B, C are intercepted by C.



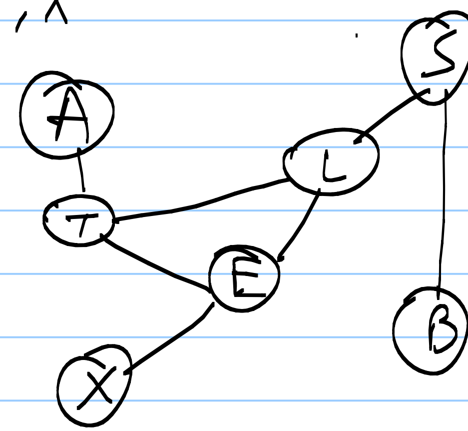


⇒ Are T and B d-separated given A and X? ($\mathcal{C} = \{A, X\}$)

get the ancestral graph of T, B, A, X ⇒



moralize ⇒



Not d-separated, because there is a path between T and B that is not intercepted by $\mathcal{C}$ (i.e., does not include any nodes in $\mathcal{C}$).

Defn. A Bayesian network (BN) consists of:

- a dag $(\mathcal{U}, \mathcal{E})$, where $\mathcal{U}$ is a set of variables and $\mathcal{E}$ is a set of directed edges
- a conditional probability table for each variable.

(simplified version of defn 2.3)

~~We require the following:~~

Let BN be a Bayesian network with variable set ("universe") $\mathcal{U}$ and let $P(\mathcal{U})$ be a probability distribution that reflects the d-separation properties specified by BN, i.e.: (1) the conditional probabilities for a variable are those specified in BN (by a table), and (2) if vars A and B are d-separated in BN by set $\mathcal{C}$, then

A and B are conditionally independent given C.

<u>Theorem 2.1</u> (the chain rule for Bayesian networks).
Let $\mathcal{U} = \{A_1, \ldots, A_n\}$. Let BN be any B. network over $\mathcal{U}$. Then BN specifies a unique prob. distribution (i.e., there exists a unique prob. distn that reflects BN) $P(\mathcal{U})$ given by the product of all conditional prob. tables of BN, i.e.,

$$P(\mathcal{U}) = \prod_{i=1}^{n} P(A_i \mid pa(A_i)),$$ where

$pa(A_i)$ are the parents of A_i and $P(\mathcal{U})$ reflects the properties of BN.