VE414 Lecture 24

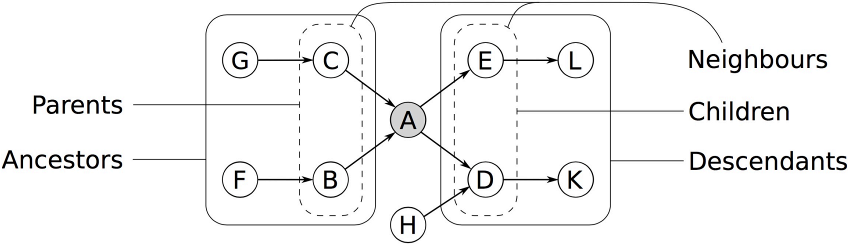
Jing Liu

UM-SJTU Joint Institute

November 28, 2019

- Recall a graph $\mathcal{G} = (\mathcal{V}, \mathcal{A})$ consists of a non-empty set $\mathcal{V}$ of nodes and a finite (but possibly empty) set $\mathcal{A}$ of paris of nodes called links.
- Each link $a = (u, v)$ can be defined either as an *ordered* or an *unordered* pair of nodes, which are said to be connected by the link.
- If (u, v) is an ordered pair, u is said to be the tail of the link and v the head, then the link is said to be directed from u to v, and is denoted by an arrow.
- The pattern with which the links appear in a graph is known as the structure.
- If all links in a graph are directed, then graph is known as a directed graph.
- The structure of a directed graph defines a partial ordering of the nodes if it does not contain any cycle or loop, such graphs are said to be acyclic.
- Causal networks (CNs) are directed graphs in which nodes represent events or variables, and links present implications. Often a causal network is acyclic.

An example of directed acyclic graph (DAG)



- A CN is a way to conduct reasoning under uncertainty given evidence.
- It is also used as a tool of forming a general model by setting up a structure:
  - Serial Connections
  - Diverging Connections
  - Converging Connections

  in which the implications are often not completely certain.

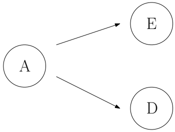
- Obviously, evidence on C will influence the certainty of A, which then alters the certainty of D. Similarly, evidence on D will alter the certainty of A, which alters the certainty of C, thus evidence can transmitted both ways.

- For example, let the rainfall intensity be C, the water level of a river nearby be A, whether having flood be D, and let them be modeled by the CN above

$$\begin{aligned} \text{Rainfall} &\in \{\text{no}, \text{light}, \text{medium}, \text{heavy}\} \\ \text{Water Level} &\in \{\text{low}, \text{medium}, \text{high}\} \\ \text{Flooding} &\in \{\text{yes}, \text{no}\} \end{aligned}$$

- However, if A is observed, then the channel is said to be blocked, and C and D become independent; no useful evidence can be transmitted.
- In this context, it means, if we already know the water level, then knowing there has been flooding will not improve our understanding about rainfall.

- The following CN, evidence on A alters the certainty of E and D, and vice versa, but evidence on E actually alters the certainty of D, and vice versa.

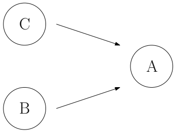


- For example, let someone's gender be A, Hair length be E, and height be D

$$\begin{aligned} \text{Gender} &\in \{\text{female}, \text{male}\} \\ \text{Hair Length} &\in \{\text{long}, \text{short}\} \\ \text{Height} &\in \{< 168\text{cm}, \geq 168\text{cm}\} \end{aligned}$$

- However, A is observed, then the channel is blocked by evidence, and E and D become independent; no more evidence can be transmitted.
- In this context, it means, if we already know it is a male, then knowing the length of his hair will not improve our understanding on his height.

Q: Lastly, consider a converging connection, are B and C independent given A?

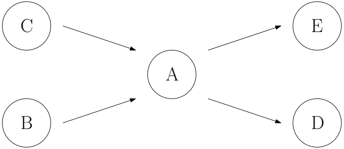


- If A is observed, the channel is said to be opened by evidence, only then B and C are not independent; evidence can be transmitted via this connection.
- For example, let whether having a broken leg be B, whether being late for this class be A, whether having a long queue in the canteen be C,

$$\begin{aligned} \text{Leg} &\in \{\text{No}, \text{Yes}\} \\ \text{Late} &\in \{\text{No}, \text{Yes}\} \\ \text{Long} &\in \{\text{No}, \text{Yes}\} \end{aligned}$$

- When we have no evidence on lateness, then the other two are independent.
- However, given being late, no having a broken leg increases the certainty of there being a long queue in the canteen, vice versa, thus dependent.

Q: Do you think having observed E or D will open the channel,



that is, make C and D dependent on each other? For example, let having a low score in my quiz be E, and let no understanding my lecture be D.

$$\text{Low} \in \{\text{No}, \text{Yes}\}$$
$$\text{lect} \in \{\text{No}, \text{Yes}\}$$

- If a variable is observed, we call it hard evidence; otherwise, it is called soft.
- Blocking in the case of serial and diverging connections requires having hard evidence, whereas opening in the case of converging connections holds for all kinds of evidence.

- The 3 types of connections cover all ways in which

  knowledge/information/evidence

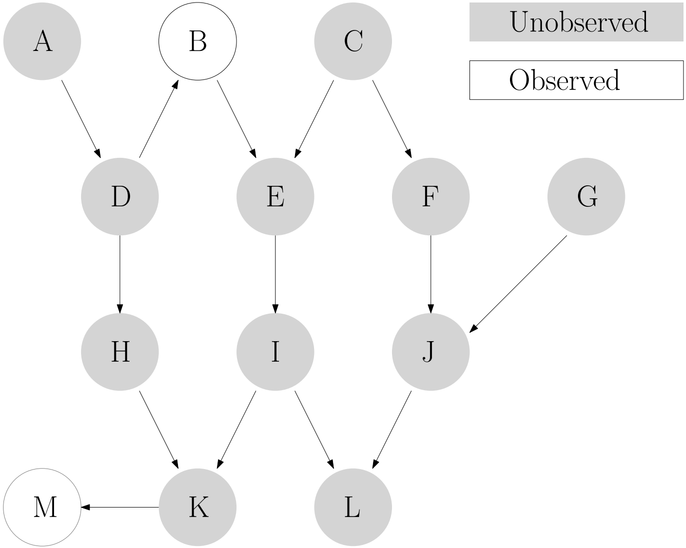
  may be transmitted through a node, thus the variable, in a CN,
  - Serial connections

    evidence may be transmitted through a serial connection when there is no hard evidence on the variable in connection.
  - Diverging connections

    evidence may be transmitted through a diverging connection when we have no hard evidence on the variable in connection.
  - Converging connections

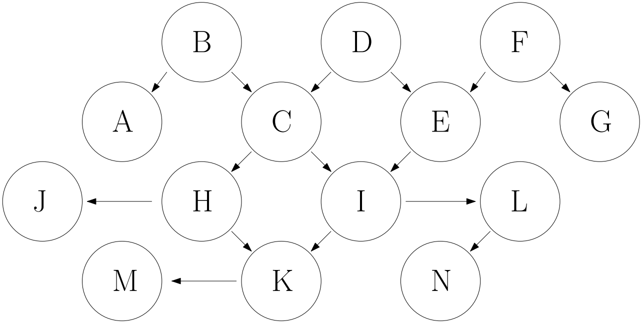
    evidence may be transmitted through a converging connection when we have evidence on the variable in connect.
- The concept of independence is known as d-separation in a graphical model.
- Given two nodes u and v in a directed graph, u and v are d-separated if all links between u and v do not transmit evidence; otherwise, are d-connected.

Q: Is A d-connected from L in the following directed graph?



- In connection to d-separation, there is a special set of nodes for a node A, known as Markov Blanket for A. It has the property that when observed, A is d-separated from the rest of the network.
- It can be shown the Markov blanket of A is the set consisting of the parents of A, the children of A, and the variables sharing a child with A.

Q: What is the Markov blanket for I in the following direct graph?



- Probabilistic graphical models (PGMs) are graphs in which nodes represent random variables, and the links represent conditional (in)dependence.
- When conditional probability is used as the "strength" of implication in *acyclic* CNs, then we have PGMs, and known as Bayesian Networks (BNs).
- So PGMs and CNs are generalisation of BNs, and CNs were used to motivate BNs. In terms of PGMs, BNs are PMGs that are directed and acyclic.
- By itself, a Bayesian network is defined a graph consisting of the followings:
- A set of random variables, and a set of directed links between them.
- Each variable has a finite set of mutually exclusive states
- The random variables together with the directed links form a DAG
- Each random variable A, with parents $B_1, \ldots, B_k$, has a conditional PMF

$$f_{A|B_1,B_2,\ldots,B_k}$$

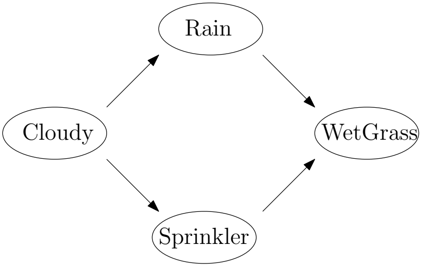
If A has no parents, then the conditional PMF reduces to a marginal PMF.

$$f_A$$

- A Bayesian network with small number of states is often presented with PMF

C	$\Pr(R=F)$	$\Pr(R=T)$
F	0.8	0.2
T	0.2	0.8

$\Pr(C=F)$	$\Pr(C=T)$
0.5	0.5



S	R	$\Pr(W=F)$	$\Pr(W=T)$
F	F	1.00	0.00
T	F	0.10	0.90
F	T	0.10	0.90
T	T	0.01	0.99

C	$\Pr(S=F)$	$\Pr(S=T)$
F	0.5	0.5
T	0.9	0.1

Q: Given the above BN, what is the chance of raining given grass is wet?

Q: What is the chance of the sprinkler is on given raining and grass is wet?

- Given a BN, it should be clear that BN allows to answers questions like:

  Is A dependent on B given that we know C?

- It should also be clear that BN is a tool to build or present a general model. It allows us to connect various components or layers of a complicated system into a a single model, and specify the strength of each implication using

$$f_{A|B_1,B_2,\ldots,B_k}$$

- Consider model the following:

Poker Game

In a certain poker game, each player receives three cards and is allowed two rounds of changing cards. In the first round, you may discard any number of cards from your hand and get replacements from the pack of cards. In the second round, you may discard at most two cards.

Q: How to construct a model to estimate my opponent's hand.