Announcements

If you are taking this class for credit, and this is your first class with me, please make an appointment to chat.

Today is filled up, but tomorrow can work.

Otherwise, next week (office hours preferable) via email.

Graphical Models

Reference for some of today's material

Chapter 8 of Bishop, available on-line

http://research.microsoft.com/~cmbishop/PRML

and Kollar and Friedman, Chapter 3.

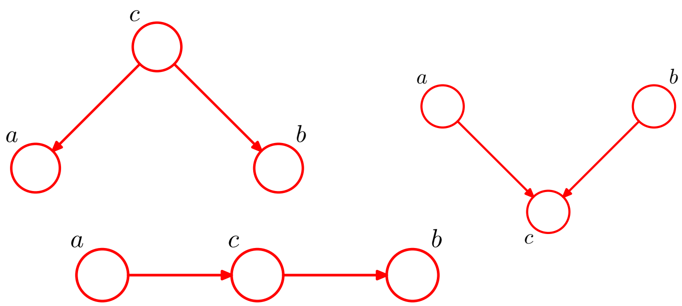
Note that Bishop uses $\perp\!\!\!\perp$ instead of $\perp$

Directed graphical models

- Graphical representation of statistical models
- DAG's (Directed acyclic graphs)
- Nodes
  - Random variables (or groups of them)
- Edges
  - Probabilistic relationships between nodes
  - In particular for directed graphical models, we have CPD (conditional probability distributions) for child given parent.
- Graphs represent factorizations of p()

Directed graphical models with 3 variables

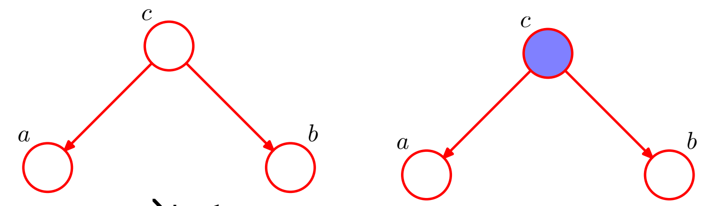
Three interesting cases



For each case, consider two questions:

1) Is $a \perp b$?
2) Is $a \perp b \mid c$? (i.e. c is observed)

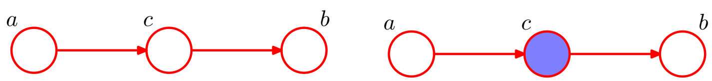
Case one (tail-to-tail) summary



$a \not\perp b$ $\quad$ $a \perp b \mid c$

Tail-to-tail case
With no conditioning, no independence
With conditioning, we have independence

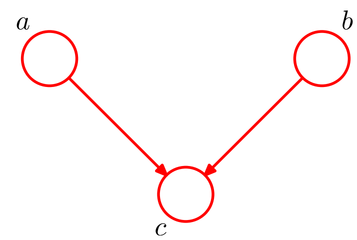
Case two (head-to-tail) summary



$a \not\perp b$ $\quad$ $a \perp b \mid c$

Head-to-tail case
With no conditioning, no independence
With conditioning, we have independence

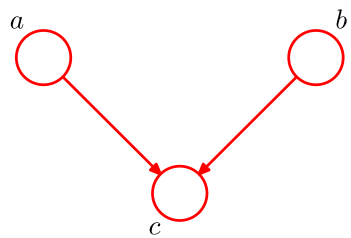
Case three



Is $a \perp b$?

Example:
c == “strange noises at night”
a == “burglar in the house”
b == “deer in the back yard”

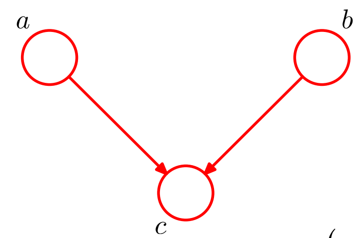
Case three



$a \perp b$

Intuitively, the arrows say that *a* is independent of *b*.

Case three

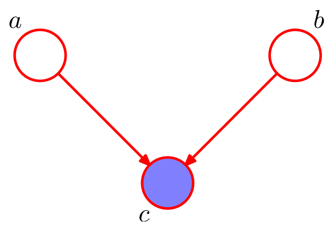


$a \perp b$

$$p(a,b,c) = p(a)\,p(b)\,p(c|a,b)$$

$$\begin{aligned} p(a,b) &= \sum_c p(a)\,p(b)\,p(c|a,b) \\ &= p(a)\,p(b)\sum_c p(c|a,b) \\ &= p(a)\,p(b) \end{aligned}$$

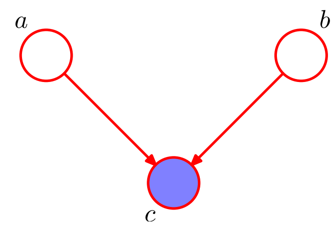
Case three with *c* observed



Is $a \perp b \mid c$?

Recall our example:
- c == "strange noises at night"
- a == "burglar in the house"
- b == "deer in the back yard"

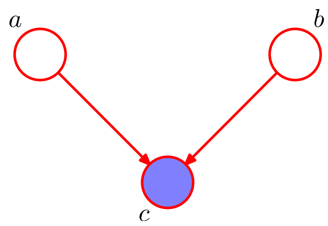
Case three with *c* observed



$a \not\perp b \mid c$

Intuitive explanation:
Given that we observe "strange noises", the two causes are anti-correlated (and thus not independent) due to"explaining away".

Case three with c observed

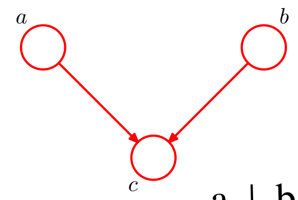


$$a \not\perp b \mid c$$

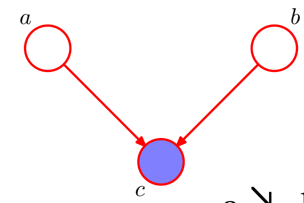
Algebraic explanation:

$$
\begin{aligned}
p(a,b|c) &= \frac{p(a,b,c)}{p(c)} \\
&= \frac{p(a)\,p(b)\,p(c|b,c)}{p(c)} \\
&= p(a)\,p(b)\frac{p(c|b,c)}{p(c)} \\
&\neq p(a)\,p(b) \qquad \text{(in general)}
\end{aligned}
$$

Case three (head-to-head) summary



$$a \perp b$$



$$a \not\perp b \mid c$$

Head-to-head case (different than the other two)
- With no conditioning, we have independence
- With conditioning, we do not have independence

If you are having trouble with "explaining away", please study Bishop, chapter 8, pages 378-379 (on-line).

Three random variables summary

In cases one and two, *a* and *b* were not independent until the observation of *c* "blocked" the (connection) path from *a* to *b*.

(From Koller and Friedman, a path that is not blocked is "active")

In case three, if *c* is not observed, the path is blocked. Observing *c* made the connection (path) active.

d-Separation (Pearl, 88)

Generalizes the examples we have been studying.

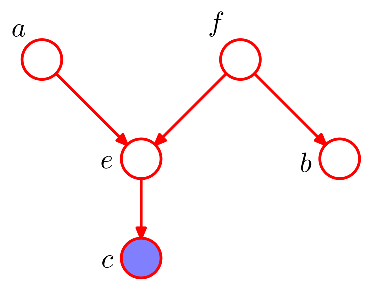
Consider non-overlapping subsets A, B, C of nodes of a graph.

Consider all paths from nodes in A to nodes in B.

A path is blocked if either:
a) The arrows meet either tail-to-tail or head-to-tail at a node in C.
b) The arrows meet head-to-head at some node that is not in C, nor
 are any of its descendants in C.

If all paths are blocked, then A and B are independent given C.

d-Separation (example one)

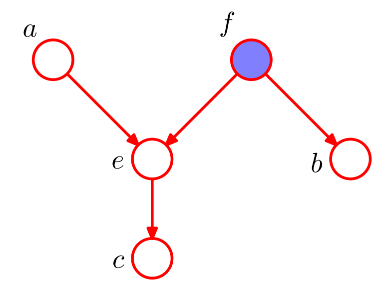


$A \not\perp B | C$

The path is not blocked by *e* because, although it is head-to-head, it has a descendant, *c*, in the conditioning set.

The path is not blocked by *f* because it is tail-to-tail, and *f* is not in C.

d-Separation (example two)



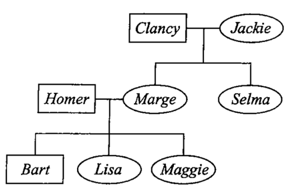
$A \perp B | F$

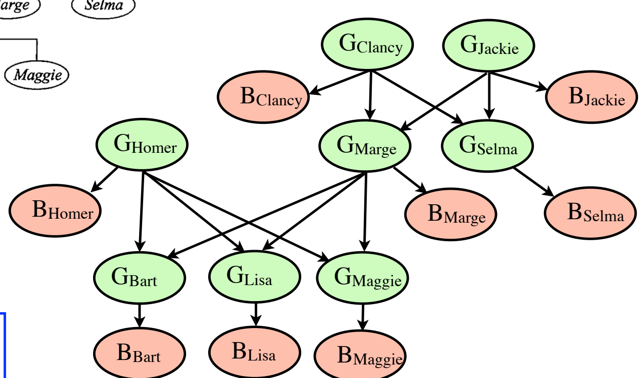
The path is blocked by *e* because it is head-to-head, and neither it, nor any of its descendants are in the conditioning set.

The path is also blocked by *f* because it is tail-to-tail, and *f* is in C.

Grounded example of a Bayesian Network

From Kollar and Friedman





G is genotype
B is blood type

Bayesian network semantics

- Represents a factorization of p()
  - Random variables are nodes
  - Factors are CPD (conditional probability distributions) for child given parent (just p(`NODE`) if no parents).

Equivalent semantic specification (Proof is in K&F, ch. 3)

- For each $X_i : X_i \perp \text{NonDescendents}(X_i) \mid \text{Parents}(X_i)$
  - Notice no mention of factorization

Conditional independence in distributions and graphs

Let $I(P)$ be the set of independence assertions of the form $(X \perp Y|Z)$ that are true for a distribution P.

Let $I(G)$ be the set of independence assertions represented by a DAG, G.

G is an I-map for P if I(G) $\subseteq$ $I(P)$

In other words, all indepence represented in G are true.
(There could be some more in P that G does nor reveal).

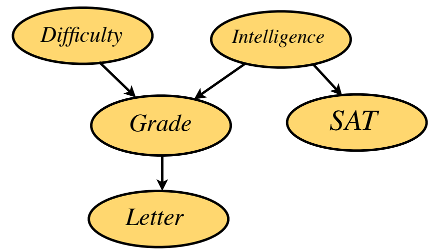
Example going from I-map to a factorization

From Kollar and Friedman

For $P(I,D,G,L,S)$, suppose $I(Graph)$ tells us

$(D \perp I|\varnothing)$ $(D \perp I|S)$ $(L \perp I,D,S|G)$ $(S \perp D,G,L|I)$ $(G \perp S|I,D)$

Note that $(A \perp B,C|X)$ just means $(A \perp B|X)$ and $(A \perp C|X)$



Example going from I-map to a factorization

For $P(I,D,G,L,S)$, suppose $I(Graph)$ tells us

$(D \perp I|\varnothing)$ $(D \perp I|S)$ $(L \perp I,D,S|G)$ $(S \perp D,G,L|I)$ $(G \perp S|I,D)$

Note that $(A \perp B,C|X)$ just means $(A \perp B|X)$ and $(A \perp C|X)$

$$P(I,D,G,L,S) = P(I)P(D|I)P(G|I,D)P(L|I,D,G)P(S|I,D,G,L)$$

Reminder about independence

$$(A \perp B|\varnothing) \Rightarrow P(A|B) = P(A)$$

$$P(A|B) = \frac{P(A,B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$$

$$(A \perp B|C) \Rightarrow P(A|B,C) = P(A|C)$$

$$P(A|B,C) = \frac{P(A,B,C)}{P(B,C)} = \frac{P(A|C)P(B|C)P(C)}{P(C)P(B|C)} = P(A|C)$$