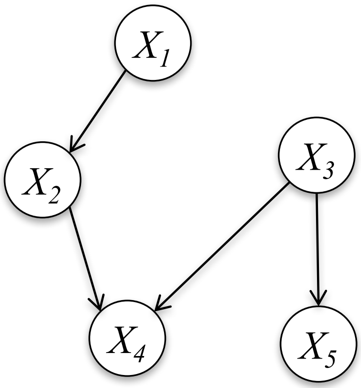
DS-GA 1018
Probabilistic Time Series Analysis
Lab 01 Bayesian Network

Ying Wang (based on slides from Matt Gormley @ CMU)

Bayesian Network

Definition:



$$P(X_1 \ldots X_n) = \prod_{i=1}^{n} P(X_i \mid parents(X_i))$$

- A Bayesian Network is a **directed graphical model**
- It consists of a graph **G** and the conditional probabilities **P**
- These two parts full specify the distribution:
  - Qualitative Specification: **G**
  - Quantitative Specification: **P**

Qualitative Specification

- Where does the qualitative specification come from?
  - Prior knowledge of causal relationships
  - Prior knowledge of modular relationships
  - Assessment from experts
  - Learning from data (i.e. structure learning)
  - We simply link a certain architecture (e.g. a layered graph)
  - …

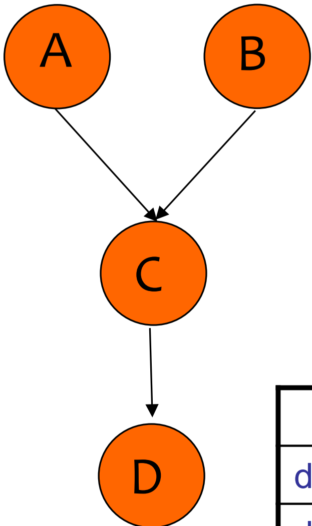
Quantitative Specification

Example: Conditional probability tables (CPTs) for discrete random variables

a^0	0.75
a^1	0.25

b^0	0.33
b^1	0.67

$$P(a,b,c.d) = P(a)P(b)P(c|a,b)P(d|c)$$



	a^0b^0	a^0b^1	a^1b^0	a^1b^1
c^0	0.45	1	0.9	0.7
c^1	0.55	0	0.1	0.3

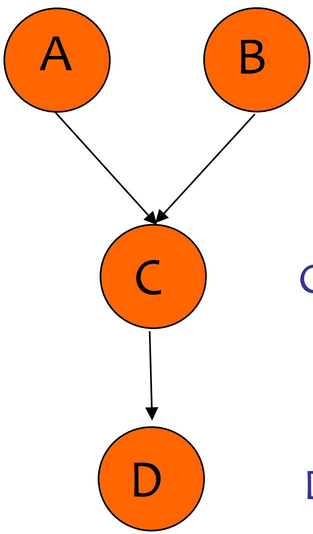
	c^0	c^1
d^0	0.3	0.5
d^1	07	0.5

Quantitative Specification

Example: Conditional probability density functions (CPDs) for continuous random variables

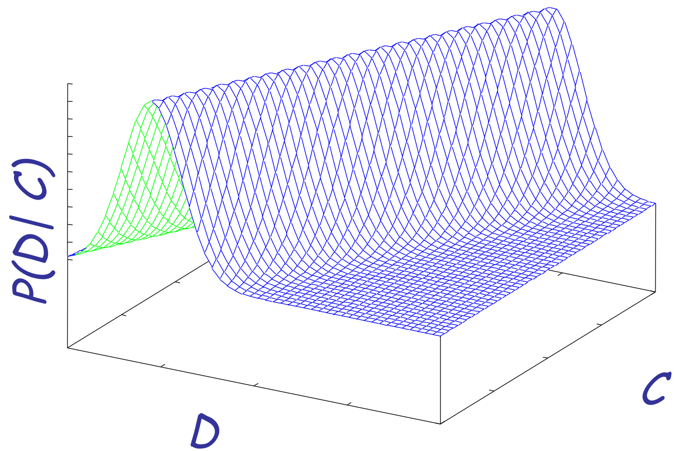
$A \sim N(\mu_a, \Sigma_a)$ $B \sim N(\mu_b, \Sigma_b)$

$$P(a,b,c.d) = P(a)P(b)P(c|a,b)P(d|c)$$



$C \sim N(A+B, \Sigma_c)$

$D \sim N(\mu_d + C, \Sigma_d)$



What Independencies does a Bayes Net Model?

- In order for a Bayesian network to model a probability distribution, the following must be true:

  Each variable is conditionally independent of all its non-descendants in the graph given the value of all its parents.

- This follows from

$$P(X_1 \ldots X_n) = \prod_{i=1}^{n} P(X_i \mid parents(X_i))$$

$$= \prod_{i=1}^{n} P(X_i \mid X_1 \ldots X_{i-1})$$

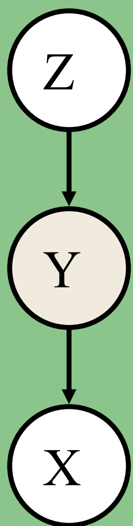
- But what else does it imply?

Slide from William Cohen

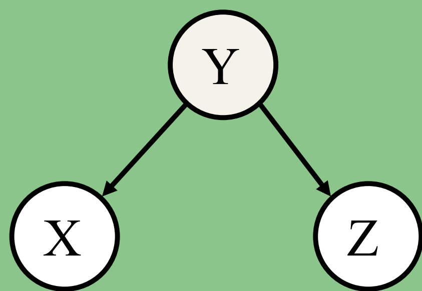
What Independencies does a Bayes Net Model?

Three cases of interest…

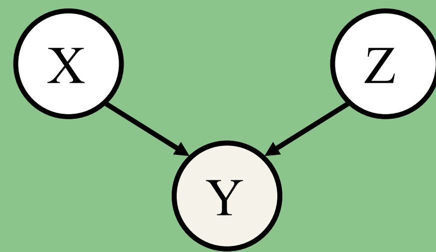
Cascade



Common Parent



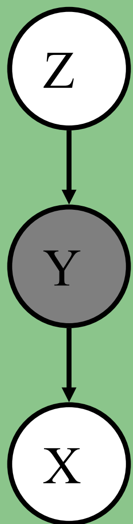
V-Structure



What Independencies does a Bayes Net Model?

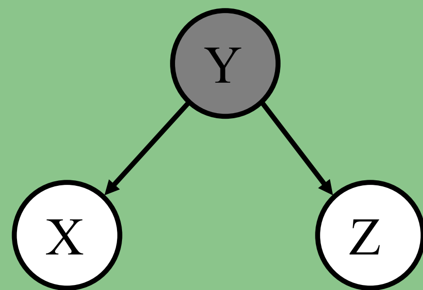
Three cases of interest…

Cascade



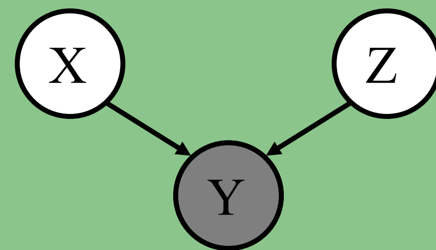
$$X \perp\!\!\!\perp Z \mid Y$$

Common Parent



$$X \perp\!\!\!\perp Z \mid Y$$

V-Structure



$$X \not\perp\!\!\!\perp Z \mid Y$$

Knowing Y
decouples X and Z

Knowing Y
couples X and Z

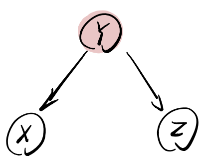
observed.

$$P(Z, X \mid Y=y) = \frac{P(Z, X, Y=y)}{P(Y=y)}$$

$$= \frac{P(Z)\,P(Y=y \mid Z)\,P(X \mid Y=y)}{P(Y=y)}$$

$$= P(Z \mid Y=y)\,P(X \mid Y=y)$$

$$\Rightarrow \quad X \perp\!\!\!\perp Z \mid Y$$

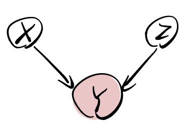


$$P(X, Z \mid Y=y) = \frac{P(X, Z, Y=y)}{P(Y=y)}$$

$$= \frac{P(X \mid Y=y)\,P(Z \mid Y=y)\,P(Y=y)}{P(Y=y)}$$

$$= P(X \mid Y=y)\,P(Z \mid Y=y)$$

$$\Rightarrow \quad X \perp\!\!\!\perp Z \mid Y$$



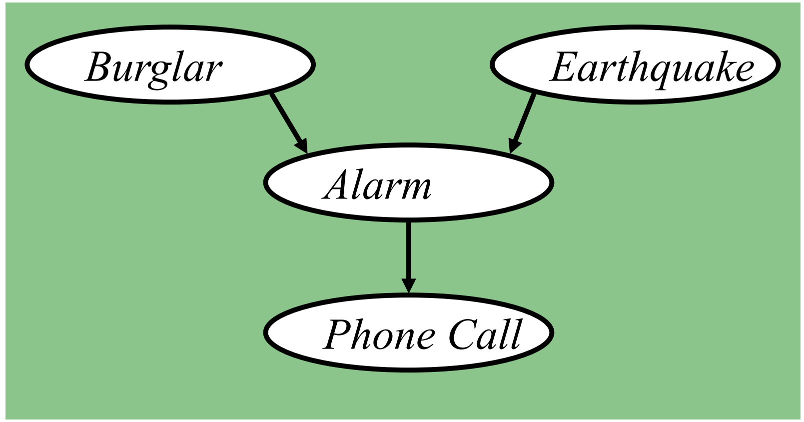
$$P(X, Z \mid Y=y) = \frac{P(X, Z, Y=y)}{P(Y=y)}$$

$$= \frac{P(X)\,P(Z)\,P(Y=y \mid X, Z)}{P(Y=y)}$$

$$\Rightarrow \quad X \not\perp\!\!\!\perp Z \mid Y$$

The "Burglar Alarm" example

- Your house has a twitchy burglar alarm that is also sometimes triggered by earthquakes.
- Earth arguably doesn't care whether your house is currently being burgled $E \perp\!\!\!\perp B$
- While you are on vacation, one of your neighbors calls and tells you your home's burglar alarm is ringing. Uh oh!



Quiz: True or False?

$$Burglar \perp\!\!\!\perp Earthquake \mid PhoneCall$$

Slide from William Cohen

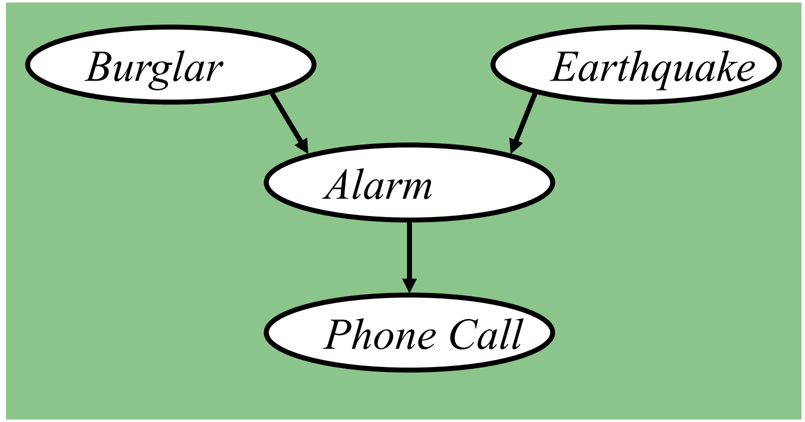
The "Burglar Alarm" example

- But now suppose you learn that there was a medium-sized earthquake in your neighborhood. Oh, whew! Probably not a burglar after all.
- Earthquake "explains away" the hypothetical burglar.
- But then it must **not** be the case that

$$Burglar \perp\!\!\!\perp Earthquake \mid PhoneCall$$

even though

$$Burglar \perp\!\!\!\perp Earthquake$$



Slide from William Cohen

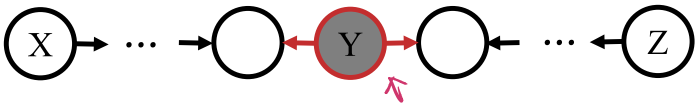
D-Separation

If variables X and Z are **d-separated** given a **set** of variables E
Then X and Z are **conditionally independent** given the **set** E

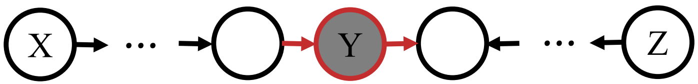
Definition #1:
Variables X and Z are **d-separated** given a **set** of evidence variables E iff every path from X to Z is "blocked".

A path is "blocked" whenever:

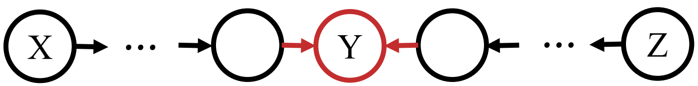
1. $\exists$ Y on path s.t. Y $\in$ E and Y is a "common parent"



2. $\exists$ Y on path s.t. Y $\in$ E and Y is in a "cascade"



3. $\exists$ Y on path s.t. {Y, descendants(Y)} $\notin$ E and Y is in a "v-structure"



Markov Blanket

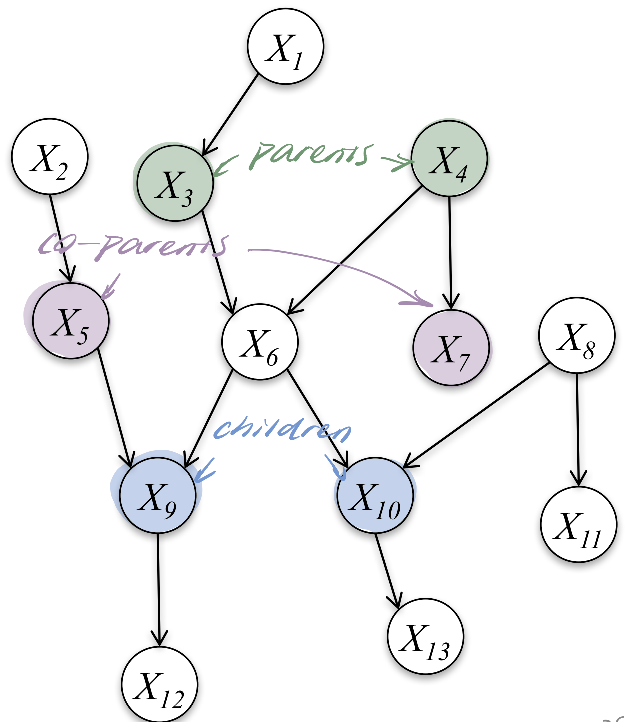
Def: the **co-parents** of a node are the parents of its children

Def: the **Markov Blanket** of a node is the set containing the node's parents, children, and co-parents.

Thm: a node is **conditionally independent** of every other node in the graph given its **Markov blanket**

Markov blanket of x_6

$\{x_3, x_4, x_5, x_8, x_9, x_{10}\}$



Consider the Boolean variables: H (it's a hot day), C (Marv is content), E (Marv eats at least one item), F (Marv catches a fish), S (Marv steals the sandwich).

a. Draw a Bayesian network for this domain. Only include the Boolean variables listed above, so your network should have 5 nodes.

$P(F|H)=x$ $\quad$ $P(F|\neg H)=y$

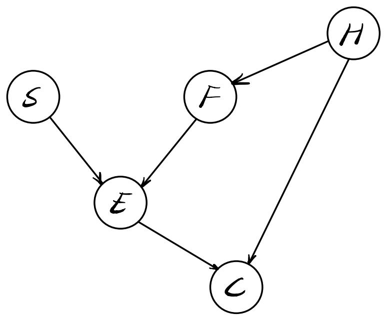
b. Suppose the probability that Marv catches the fish is *x* when it's hot, and *y* when it is not. Give the conditional probability table associated with F.

$P(E|F)=1$ $\quad$ $P(E|S)=0.5$

c. Suppose that if Marv catches a fish, he will eat it with probability 1, and if he successfully steals the sandwich, he will eat it with probability 0.5. If he fails at both hunting and stealing, then he will not eat anything. Give the conditional probability table associated with E. $P(E|\neg F,\neg S)=0$

d. Suppose Marv is content. Write down the expression for the probability he caught a fish, in terms of the various conditional probabilities in the network.

(a)



(b)

$P(F\|H)$		
	$F=1$	$F=0$
$H=1$	x	$1-x$
$H=0$	y	$1-y$

(c)

$P(E\|F,S)$		
	$E=1$	$E=0$
$F=1, S=1$	1	0
$F=1, S=0$	1	0
$F=0, S=1$	0.5	0.5
$F=0, S=0$	0	1

(d) $$P(F|C)=\frac{P(F,C)}{P(C)}$$

$$P(C)=\sum_{s,e,f,h} P(S=s, E=e, F=f, H=h, C=1)$$

$$=\sum_{s,e,f,h} P(S=s)P(E=e|S=s,F=f)P(F=f|H=h)P(H=h)P(C=1|E=e,H=h)$$

$$P(F,C)=\sum_{s,e,h} P(S=s, E=e, F=1, H=h, C=1)$$

$$=\sum_{s,e,h} P(S=s)P(E=e|S=s,F=1)P(F=1|H=h)P(H=h)P(C=1|E=e,H=h)$$

$$\Rightarrow P(F|C)=\frac{\sum_{s,e,h} P(S=s)P(E=e|S=s,F=1)P(F=1|H=h)P(H=h)P(C=1|E=e,H=h)}{\sum_{s,e,f,h} P(S=s)P(E=e|S=s,F=f)P(F=f|H=h)P(H=h)P(C=1|E=e,H=h)}$$

$s\in\{0,1\},\ e\in\{0,1\},\ f\in\{0,1\},\ h\in\{0,1\}$