CS 188: Artificial Intelligence
Spring 2006

Lecture 16: Bayes' Nets II
3/14/2006

Dan Klein – UC Berkeley

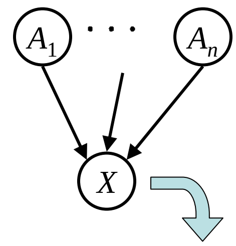
Today

- Last time: Bayes' nets
  - Introduction
  - Semantics (BN to joint distribution)
- Today:
  - Conditional independence
  - How independence determines structure
  - How structure determines independence

Bayes' Net Semantics

- A Bayes' net:
  - A set of nodes, one per variable X
  - A directed, acyclic graph
  - A conditional distribution of each variable conditioned on its parents

$$P(X|a_1 \dots a_n)$$

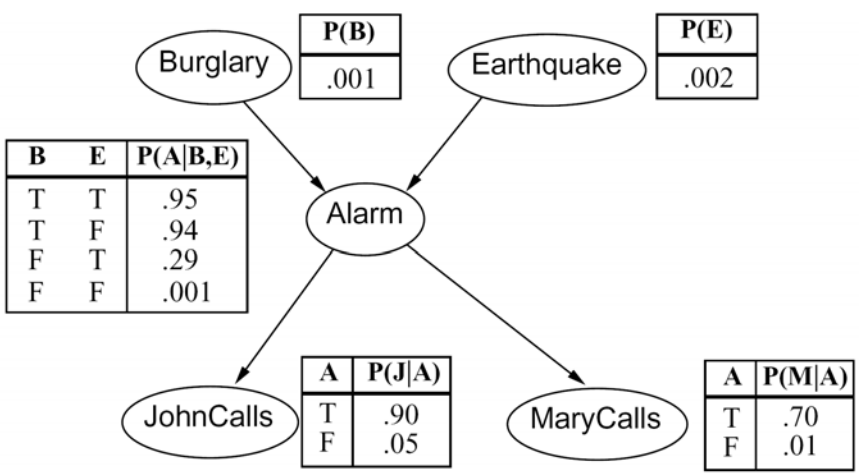


$$P(X|A_1 \dots A_n)$$

- Semantics:
  - A BN defines a joint probability distribution over its variables:

$$P(x_1, x_2, \dots x_n) = \prod_{i=1}^{n} P(x_i|\textit{parents}(X_i))$$

Example: Alarm Network



P(B)
.001

P(E)
.002

B	E	P(A|B,E)
T	T	.95
T	F	.94
F	T	.29
F	F	.001

A	P(J|A)
T	.90
F	.05

A	P(M|A)
T	.70
F	.01

$$P(b, e, \neg a, j, m) =$$

Bayes' Nets

- So far, we talked about how a Bayes' net encodes a joint distribution
- Next: how to answer queries about that distribution
  - Key ingredient: conditional independence
  - Last class: assembled BNs using an intuitive notion of conditional independence as causality
  - Today: formalize these ideas
  - Main goal: answer queries about conditional independence and influence
- After that: how to answer numerical queries (inference)

Conditional Independence

- Reminder: independence
  - X and Y are independent if

$$\forall x, y \;\; P(x, y) = P(x)P(y) \dashrightarrow X \perp\!\!\!\perp Y$$

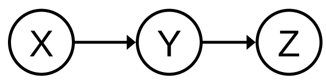
  - X and Y are conditionally independent given Z

$$\forall x, y, z \;\; P(x, y|z) = P(x)P(y) \dashrightarrow X \perp\!\!\!\perp Y|Z$$

  - (Conditional) independence is a property of a distribution

Independence in a BN

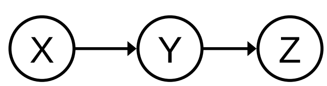
- Important question about a BN:
  - Are two nodes independent given certain evidence
  - If yes, can calculate using algebra (really tedious)
  - If no, can prove with a counter example
  - Example:



  - Question: are X and Z independent?
    - Answer: not *necessarily*, we've seen examples otherwise: low pressure causes rain which causes traffic.
    - X can influence Z, Z can influence X
    - Addendum: they *could* be independent: how?

Causal Chains

- This configuration is a "causal chain"



X: Low pressure

Y: Rain

Z: Traffic

$$P(x, y, z) = P(x)P(y|x)P(z|y)$$

  - Is X independent of Z given Y?

$$P(z|x, y) = \frac{P(x, y, z)}{P(x, y)} = \frac{P(x)P(y|x)P(z|y)}{P(x)P(y|x)}$$

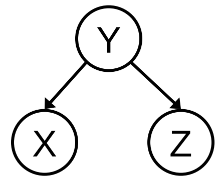
$$= P(z|y)$$

Yes!

  - Evidence along the chain "blocks" the influence

Common Cause

- Another basic configuration: two effects of the same cause
  - Are X and Z independent?
    - No, remember the “project due” example
  - Are X and Z independent given Y?



Y: Project due

X: Newsgroup busy

Z: Lab full

$$P(z|x,y) = \frac{P(x,y,z)}{P(x,y)} = \frac{P(y)P(x|y)P(z|y)}{P(y)P(x|y)}$$

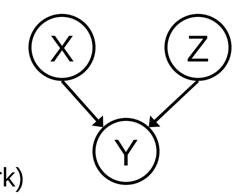
$$= P(z|y)$$

Yes!

  - Observing the cause blocks influence between effects.

Common Effect

- Last configuration: two causes of one effect (v-structures)
  - Are X and Z independent?
    - Yes: remember the ballgame and the rain causing traffic, no correlation?
    - Still need to prove they must be (homework)
  - Are X and Z independent given Y?
    - No: remember that seeing traffic put the rain and the ballgame in competition?
  - This is backwards from the other cases
  - Observing the effect enables influence between effects.



X: Raining

Z: Ballgame

Y: Traffic

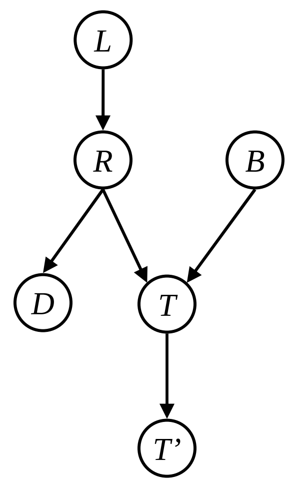
The General Case

- Any complex example can be analyzed using these three canonical cases
- General question: in a given BN, are two variables independent (given evidence)?
- Solution: graph search!

Example

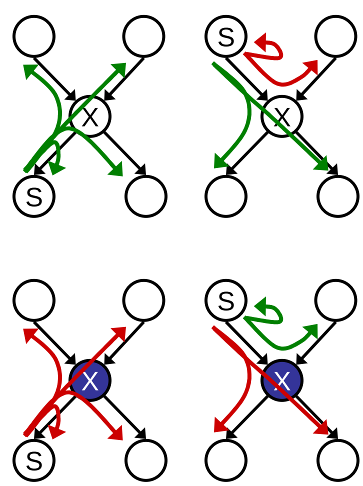
Reachability

- Recipe: shade evidence nodes
- Attempt 1: if two nodes are connected by an undirected path not blocked by a shaded node, they are conditionally independent
- Almost works, but not quite
  - Where does it break?
  - Answer: the v-structure at T doesn't count as a link in a path unless shaded



Reachability (the Bayes' ball)

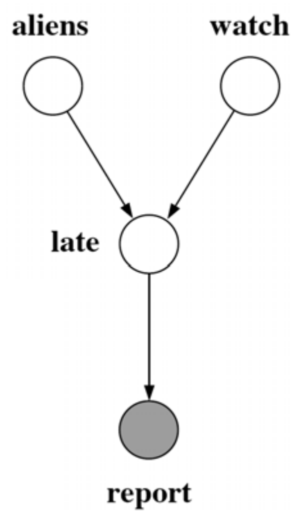
- Correct algorithm:
  - Start at source node
  - Try to reach target with graph search
  - States: node along with previous arc
  - Successor function:
    - Unobserved nodes:
      - To any child
      - To any parent if coming from a child
    - Observed nodes:
      - From parent to parent
  - If you can't reach a node, it's conditionally independent



Example

$A \perp\!\!\!\perp W$ Yes

$A \perp\!\!\!\perp W | R$



Example

- Questions:

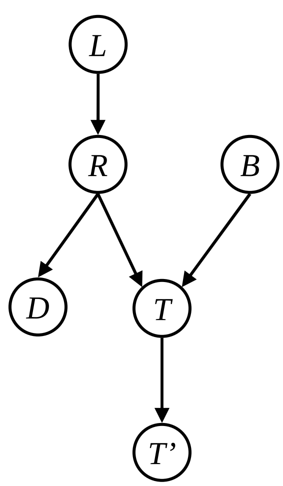
  $L \perp\!\!\!\perp T' | T$ Yes

  $L \perp\!\!\!\perp B$ Yes

  $L \perp\!\!\!\perp B | T$

  $L \perp\!\!\!\perp B | T'$

  $L \perp\!\!\!\perp B | T, R$ Yes



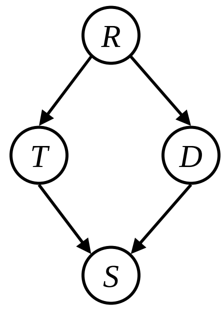
Example

- Variables:
  - R: Raining
  - T: Traffic
  - D: Roof drips
  - S: I'm sad
- Questions:

  $T \perp\!\!\!\perp D$

  $T \perp\!\!\!\perp D | R$ *Yes*

  $T \perp\!\!\!\perp D | R, S$

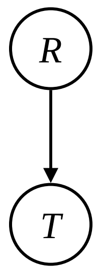


Causality?

- When Bayes' nets reflect the true causal patterns:
  - Often simpler (nodes have fewer parents)
  - Often easier to think about
  - Often easier to elicit from experts
- BNs need not actually be causal
  - Sometimes no causal net exists over the domain
  - E.g. consider the variables *Traffic* and *Drips*
  - End up with arrows that reflect correlation, not causation
- What do the arrows really mean?
  - Topology may happen to encode causal structure
  - Topology only guaranteed to encode conditional independencies

Example: Traffic

- Basic traffic net
- Let's multiply out the joint



$P(R)$

r	1/4
¬r	3/4

$P(T|R)$

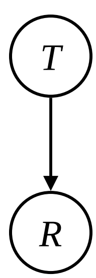
r	t	3/4
	¬t	1/4
¬r	t	1/2
	¬t	1/2

$P(T, R)$

r	t	3/16
r	¬t	1/16
¬r	t	6/16
¬r	¬t	6/16

Example: Reverse Traffic

- Reverse causality?



$P(T)$

t	9/16
¬t	7/16

$P(R|T)$

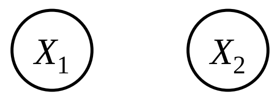
t	r	1/3
	¬r	2/3
¬t	r	1/7
	¬r	6/7

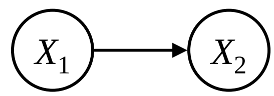
$P(T, R)$

r	t	3/16
r	¬t	1/16
¬r	t	6/16
¬r	¬t	6/16

Example: Coins

- Extra arcs don't prevent representing independence, just allow non-independence





$P(X_1)$

h	0.5
t	0.5

$P(X_2)$

h	0.5
t	0.5

$P(X_1)$

h	0.5
t	0.5

$P(X_2|X_1)$

h | h	0.5
t | h	0.5

h | t	0.5
t | t	0.5

Summary

- Bayes nets compactly encode joint distributions
- Guaranteed independencies of distributions can be deduced from BN graph structure
- The Bayes' ball algorithm (aka d-separation)
- A Bayes net may have other independencies that are not detectable until you inspect its specific distribution