I could restructure the program's flow, or use one little 'goto' instead.

Eh, screw good practice. How bad can it be?

goto main_sub3;

*Compile*

Dinosaur rider.
CS 473: Artificial Intelligence

Bayes’ Nets: Independence

Steve Tanimoto

[These slides were created by Dan Klein and Pieter Abbeel for CS188 Intro to AI at UC Berkeley. All CS188 materials are available at http://ai.berkeley.edu.]
Recap: Bayes’ Nets

- A Bayes’ net is an efficient encoding of a probabilistic model of a domain

- Questions we can ask:
  - Inference: given a fixed BN, what is $P(X \mid e)$?
  - Representation: given a BN graph, what kinds of distributions can it encode?
  - Modeling: what BN is most appropriate for a given domain?
Bayes’ Nets

- Representation
  - Conditional Independences
  - Probabilistic Inference
  - Learning Bayes’ Nets from Data
Conditional Independence

- X and Y are independent if

\[ \forall x, y \quad P(x, y) = P(x)P(y) \quad \longrightarrow \quad X \independent Y \]

- X and Y are conditionally independent given Z

\[ \forall x, y, z \quad P(x, y | z) = P(x | z)P(y | z) \quad \longrightarrow \quad X \independent Y | Z \]

- (Conditional) independence is a property of a distribution

- Example: \( \text{Alarm} \independent \text{Fire} | \text{Smoke} \)
Bayes Nets: Assumptions

- Assumptions we are required to make to define the Bayes net when given the graph:
  \[
P(x_i|x_1 \cdots x_{i-1}) = P(x_i|\text{parents}(X_i))
  \]

- Beyond above “chain rule $\to$ Bayes net” conditional independence assumptions
  - Often additional conditional independences
  - They can be read off the graph

- Important for modeling: understand assumptions made when choosing a Bayes net graph
Independence in a BN

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


![Diagram of nodes X, Y, Z connected in a chain]

- Question: are X and Z necessarily independent?
  - Answer: no. Example: low pressure causes rain, which causes traffic.
  - X can influence Z, Z can influence X (via Y)
  - Addendum: they could be independent: how?
D-separation: Outline
D-separation: Outline

- Study independence properties for triples
- Analyze complex cases in terms of member triples
- D-separation: a condition / algorithm for answering such queries
Causal Chains

- This configuration is a “causal chain”

- Guaranteed X independent of Z?  **No!**

- One example set of CPTs for which X is not independent of Z is sufficient to show this independence is not guaranteed.

- Example:
  - Low pressure causes rain causes traffic, high pressure causes no rain causes no traffic

- In numbers:
  \[
P(+y \mid +x) = 1, \quad P(-y \mid -x) = 1, \\
P(+z \mid +y) = 1, \quad P(-z \mid -y) = 1
\]
Causal Chains

- This configuration is a “causal chain”

\[
P(x, y, z) = P(x)P(y|x)P(z|y)
\]

- Guaranteed 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

- This configuration is a “common cause”

- Guaranteed X independent of Z? **No!**
  - One example set of CPTs for which X is not independent of Z is sufficient to show this independence is not guaranteed.
  - Example:
    - Project due causes both forums busy and lab full
    - In numbers:
      \[
      P( +x \mid +y ) = 1, \quad P( -x \mid -y ) = 1,
      \]
      \[
      P( +z \mid +y ) = 1, \quad P( -z \mid -y ) = 1
      \]

\[
P(x, y, z) = P(y)P(x\mid y)P(z\mid y)
\]
Common Cause

- This configuration is a “common cause”

\[ P(x, y, z) = P(y)P(x|y)P(z|y) \]

- Guaranteed X and Z independent given Y?

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

Yes!

- Observing the cause blocks influence between effects.
Common Effect

- Last configuration: two causes of one effect (v-structures)

- Are X and Y independent?
  - Yes: the ballgame and the rain cause traffic, but they are not correlated
  - Still need to prove they must be (try it!)

- Are X and Y independent given Z?
  - No: seeing traffic puts the rain and the ballgame in competition as explanation.

- This is backwards from the other cases
  - Observing an effect activates influence between possible causes.
The General Case
The General Case

- General question: in a given BN, are two variables independent (given evidence)?
- Solution: analyze the graph

- Any complex example can be broken into repetitions of the three canonical cases
Reachability

- Recipe: shade evidence nodes, look for paths in the resulting graph

- Attempt 1: if two nodes are connected by an undirected path not blocked by a shaded node, then they are not 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 “active”
**Active / Inactive Paths**

- **Question:** Are X and Y conditionally independent given evidence variables \( \{Z\} \)?
  - Yes, if X and Y “d-separated” by Z
  - Consider all (undirected) paths from X to Y
  - No active paths = independence!

- **A path is active if each triple is active:**
  - Causal chain \( A \rightarrow B \rightarrow C \) where B is unobserved (either direction)
  - Common cause \( A \leftarrow B \rightarrow C \) where B is unobserved
  - Common effect (aka v-structure)
    - \( A \rightarrow B \leftarrow C \) where B or one of its descendents is observed

- **All it takes to block a path is a single inactive segment**
D-Separation

- **Query:** \( X_i \perp \perp X_j \mid \{X_{k_1}, \ldots, X_{k_n}\} \) ?

- **Check all (undirected!) paths between** \( X_i \) and \( X_j \)
  - If one or more active, then independence not guaranteed
    \[
    X_i \nmid X_j \mid \{X_{k_1}, \ldots, X_{k_n}\}
    \]
  - Otherwise (i.e. if all paths are inactive),
    then independence is guaranteed
    \[
    X_i \perp \perp X_j \mid \{X_{k_1}, \ldots, X_{k_n}\}
    \]
Example

\[ R \perp B \quad \text{Yes} \]
\[ R \perp B \mid T \]
\[ R \perp B \mid T' \]
Example

$L \perp T' | T$  Yes
$L \perp B$  Yes
$L \perp B | T$
$L \perp B | T'$
$L \perp B | T, R$  Yes
Example

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

- Questions:

\[
T \perp D \\
T \perp D|R \\
T \perp D|R, S
\]

Yes
Structure Implications

- Given a Bayes net structure, can run d-separation algorithm to build a complete list of conditional independences that are necessarily true of the form

\[ X_i \perp\!
\!
\perp X_j \mid \{X_{k_1}, \ldots, X_{k_n}\} \]

- This list determines the set of probability distributions that can be represented
Computing All Independences!
Topology Limits Distributions

- Given some graph topology $G$, only certain joint distributions can be encoded.
- The graph structure guarantees certain (conditional) independences.
- (There might be more independence)
- Adding arcs increases the set of distributions, but has several costs.
- Full conditioning can encode any distribution.

\[
\{X \perp Y, X \perp Z, Y \perp Z, \\
X \perp Z \mid Y, X \perp Y \mid Z, Y \perp Z \mid X\}
\]
Bayes Nets Representation Summary

- Bayes nets compactly encode joint distributions
- Guaranteed independencies of distributions can be deduced from BN graph structure
- D-separation gives precise conditional independence guarantees from graph alone
- A Bayes’ net’s joint distribution may have further (conditional) independence that is not detectable until you inspect its specific distribution
Bayes’ Nets

- Representation
- Conditional Independences
- Probabilistic Inference
  - Enumeration (exact, exponential complexity)
  - Variable elimination (exact, worst-case exponential complexity, often better)
  - Probabilistic inference is NP-complete
  - Sampling (approximate)

- Learning Bayes’ Nets from Data