Bayesian Networks

Vibhav Gogate
The University of Texas at Dallas

Intro to AI (CS 4365)

Many slides over the course adapted from either Dan Klein, Luke Zettlemoyer, Stuart Russell or Andrew Moore
Outline

- Probabilistic models (and inference)
  - Bayesian Networks (BNs)
  - Independence in BNs
Bayes’ Nets: Big Picture

- Two problems with using full joint distribution tables as our probabilistic models:
  - Unless there are only a few variables, the joint is WAY too big to represent explicitly
  - Hard to learn (estimate) anything empirically about more than a few variables at a time

- **Bayes’ nets:** a technique for describing complex joint distributions (models) using simple, local distributions (conditional probabilities)
  - More properly called *graphical models*
  - We describe how variables locally interact
  - Local interactions chain together to give global, indirect interactions
Bayes’ Net Semantics

- Let’s formalize the semantics of a Bayes’ net
- A set of nodes, one per variable X
- A directed, acyclic graph
- A conditional distribution for each node
  - A collection of distributions over X, one for each combination of parents’ values
    \[
    P(X|a_1 \ldots a_n)
    \]
- CPT: conditional probability table

A Bayes net = Topology (graph) + Local Conditional Probabilities
Example Bayes’ Net: Car
Probabilities in BNs

- Bayes’ nets implicitly encode joint distributions
  - As a product of local conditional distributions
  - To see what probability a BN gives to a full assignment, multiply all the relevant conditionals together:

\[
P(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} P(x_i | \text{parents}(X_i))
\]

- This lets us reconstruct any entry of the full joint

- Not every BN can represent every joint distribution
  - The topology enforces certain independence assumptions
  - Compare to the exact decomposition according to the chain rule!
Example Bayes’ Net: Insurance
Example: Independence

- N fair, independent coin flips:

\[
P(X_1) \quad P(X_2) \quad \ldots \quad P(X_n)
\]

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>h</td>
<td>0.5</td>
<td>h</td>
<td>0.5</td>
</tr>
<tr>
<td>t</td>
<td>0.5</td>
<td>t</td>
<td>0.5</td>
</tr>
</tbody>
</table>

\[P(X_1, X_2, \ldots, X_n)\]

\[2^n\]
Example: Coin Flips

- N independent coin flips

\[ X_1 \quad X_2 \quad \ldots \quad X_n \]

- No interactions between variables: absolute independence
Independence

- Two variables are *independent* if:

\[ \forall x, y : P(x, y) = P(x)P(y) \]

- This says that their joint distribution *factors* into a product two simpler distributions
- Another form:

\[ \forall x, y : P(x|y) = P(x) \]

- We write: \( X \perp\!\!\!\!\!\!\perp Y \)
Independence

- Two variables are *independent* if:

  \[ \forall x, y : P(x, y) = P(x)P(y) \]

  - This says that their joint distribution *factors* into a product two simpler distributions
  - Another form:

  \[ \forall x, y : P(x|y) = P(x) \]

  - We write: \( X \perp\!\!\!\!\!\!\!\perp Y \)

- Independence is a simplifying *modeling assumption*
  - *Empirical* joint distributions: at best “close” to independent
  - What could we assume for \{Weather, Traffic, Cavity, Toothache\}?
Example: Independence?

\[ P_1(T, W) \]

<table>
<thead>
<tr>
<th>T</th>
<th>W</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>warm</td>
<td>sun</td>
<td>0.4</td>
</tr>
<tr>
<td>warm</td>
<td>rain</td>
<td>0.1</td>
</tr>
<tr>
<td>cold</td>
<td>sun</td>
<td>0.2</td>
</tr>
<tr>
<td>cold</td>
<td>rain</td>
<td>0.3</td>
</tr>
</tbody>
</table>

\[ P_2(T, W) \]

<table>
<thead>
<tr>
<th>T</th>
<th>W</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>warm</td>
<td>sun</td>
<td>0.3</td>
</tr>
<tr>
<td>warm</td>
<td>rain</td>
<td>0.2</td>
</tr>
<tr>
<td>cold</td>
<td>sun</td>
<td>0.3</td>
</tr>
<tr>
<td>cold</td>
<td>rain</td>
<td>0.2</td>
</tr>
</tbody>
</table>
**Example: Independence?**

\[
P_1(T, W) \\
\begin{array}{|c|c|c|}
\hline
T & W & P \\
\hline
\text{warm} & \text{sun} & 0.4 \\
\text{warm} & \text{rain} & 0.1 \\
\text{cold} & \text{sun} & 0.2 \\
\text{cold} & \text{rain} & 0.3 \\
\hline
\end{array}
\]

\[
P_2(T, W) \\
\begin{array}{|c|c|c|}
\hline
T & P \\
\hline
\text{warm} & 0.5 \\
\text{cold} & 0.5 \\
\hline
\end{array}
\]

\[
P_2(T, W) \\
\begin{array}{|c|c|c|}
\hline
T & W & P \\
\hline
\text{warm} & \text{sun} & 0.3 \\
\text{warm} & \text{rain} & 0.2 \\
\text{cold} & \text{sun} & 0.3 \\
\text{cold} & \text{rain} & 0.2 \\
\hline
\end{array}
\]
Example: Independence?

\[ P_1(T, W) \]

<table>
<thead>
<tr>
<th>T</th>
<th>W</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>warm</td>
<td>sun</td>
<td>0.4</td>
</tr>
<tr>
<td>warm</td>
<td>rain</td>
<td>0.1</td>
</tr>
<tr>
<td>cold</td>
<td>sun</td>
<td>0.2</td>
</tr>
<tr>
<td>cold</td>
<td>rain</td>
<td>0.3</td>
</tr>
</tbody>
</table>

\[ P_2(T, W) \]

<table>
<thead>
<tr>
<th>T</th>
<th>W</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>warm</td>
<td>sun</td>
<td>0.3</td>
</tr>
<tr>
<td>warm</td>
<td>rain</td>
<td>0.2</td>
</tr>
<tr>
<td>cold</td>
<td>sun</td>
<td>0.3</td>
</tr>
<tr>
<td>cold</td>
<td>rain</td>
<td>0.2</td>
</tr>
</tbody>
</table>

\[ P(T, P) \]

<table>
<thead>
<tr>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>warm</td>
<td>0.5</td>
</tr>
<tr>
<td>cold</td>
<td>0.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>W</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>sun</td>
<td>0.6</td>
</tr>
<tr>
<td>rain</td>
<td>0.4</td>
</tr>
</tbody>
</table>
Example: Independence?

\[
P_1(T, W)
\]

<table>
<thead>
<tr>
<th>T</th>
<th>W</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>warm</td>
<td>sun</td>
<td>0.4</td>
</tr>
<tr>
<td>warm</td>
<td>rain</td>
<td>0.1</td>
</tr>
<tr>
<td>cold</td>
<td>sun</td>
<td>0.2</td>
</tr>
<tr>
<td>cold</td>
<td>rain</td>
<td>0.3</td>
</tr>
</tbody>
</table>

\[
P_2(T, W)
\]

<table>
<thead>
<tr>
<th>T</th>
<th>W</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>warm</td>
<td>sun</td>
<td>0.3</td>
</tr>
<tr>
<td>warm</td>
<td>rain</td>
<td>0.2</td>
</tr>
<tr>
<td>cold</td>
<td>sun</td>
<td>0.3</td>
</tr>
<tr>
<td>cold</td>
<td>rain</td>
<td>0.2</td>
</tr>
</tbody>
</table>

\[
P(T)
\]

<table>
<thead>
<tr>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>warm</td>
<td>0.5</td>
</tr>
<tr>
<td>cold</td>
<td>0.5</td>
</tr>
</tbody>
</table>

\[
P(W)
\]

<table>
<thead>
<tr>
<th>W</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>sun</td>
<td>0.6</td>
</tr>
<tr>
<td>rain</td>
<td>0.4</td>
</tr>
</tbody>
</table>
Conditional Independence

- \( P(\text{Toothache, Cavity, Catch}) \)
- If I have a cavity, the probability that the probe catches in it doesn't depend on whether I have a toothache:
  - \( P(+\text{catch} | +\text{toothache, } +\text{cavity}) = P(+\text{catch} | +\text{cavity}) \)
- The same independence holds if I don’t have a cavity:
  - \( P(+\text{catch} | +\text{toothache, } -\text{cavity}) = P(+\text{catch} | -\text{cavity}) \)
- Catch is \textit{conditionally independent} of Toothache given Cavity:
  - \( P(\text{Catch} | \text{Toothache, Cavity}) = P(\text{Catch} | \text{Cavity}) \)
- Equivalent statements:
  - \( P(\text{Toothache} | \text{Catch , Cavity}) = P(\text{Toothache} | \text{Cavity}) \)
  - \( P(\text{Toothache, Catch} | \text{Cavity}) = P(\text{Toothache} | \text{Cavity}) P(\text{Catch} | \text{Cavity}) \)
  - One can be derived from the other easily
Conditional Independence

- Unconditional (absolute) independence very rare (why?)

- Conditional independence is our most basic and robust form of knowledge about uncertain environments:

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

- What about this domain:
  - Traffic
  - Umbrella
  - Raining

- What about fire, smoke, alarm?
Ghostbusters Chain Rule

- 2-position maze, each sensor indicates ghost location
- T: Top square is red
  B: Bottom square is red
  G: Ghost is in the top

<table>
<thead>
<tr>
<th>T</th>
<th>B</th>
<th>G</th>
<th>P(T,B,G)</th>
</tr>
</thead>
<tbody>
<tr>
<td>+t</td>
<td>+b</td>
<td>+g</td>
<td>0.16</td>
</tr>
<tr>
<td>+t</td>
<td>+b</td>
<td>←g</td>
<td>0.16</td>
</tr>
<tr>
<td>+t</td>
<td>←b</td>
<td>+g</td>
<td>0.24</td>
</tr>
<tr>
<td>+t</td>
<td>←b</td>
<td>←g</td>
<td>0.04</td>
</tr>
<tr>
<td>←t</td>
<td>+b</td>
<td>+g</td>
<td>0.04</td>
</tr>
<tr>
<td>←t</td>
<td>+b</td>
<td>←g</td>
<td>0.24</td>
</tr>
<tr>
<td>←t</td>
<td>←b</td>
<td>+g</td>
<td>0.06</td>
</tr>
<tr>
<td>←t</td>
<td>←b</td>
<td>←g</td>
<td>0.06</td>
</tr>
</tbody>
</table>
Ghostbusters Chain Rule

- 2-position maze, each sensor indicates ghost location
- T: Top square is red
  B: Bottom square is red
  G: Ghost is in the top
- That means, the two sensors are conditionally independent, given the ghost position
- Can assume:
  \[ P(+g) = 0.5 \]
  \[ P(+t | +g) = 0.8 \]
  \[ P(+t | \bar{g}) = 0.4 \]
  \[ P(+b | +g) = 0.4 \]
  \[ P(+b | \bar{g}) = 0.8 \]

\[
P(T,B,G) = P(G) \cdot P(T|G) \cdot P(B|G)
\]

<table>
<thead>
<tr>
<th></th>
<th>T</th>
<th>B</th>
<th>G</th>
<th>P(T,B,G)</th>
</tr>
</thead>
<tbody>
<tr>
<td>+t</td>
<td>+b</td>
<td>+g</td>
<td></td>
<td>0.16</td>
</tr>
<tr>
<td>+t</td>
<td>+b</td>
<td>\bar{g}</td>
<td></td>
<td>0.16</td>
</tr>
<tr>
<td>+t</td>
<td>\bar{b}</td>
<td>+g</td>
<td></td>
<td>0.24</td>
</tr>
<tr>
<td>+t</td>
<td>\bar{b}</td>
<td>\bar{g}</td>
<td></td>
<td>0.04</td>
</tr>
<tr>
<td>\bar{t}</td>
<td>+b</td>
<td>+g</td>
<td></td>
<td>0.04</td>
</tr>
<tr>
<td>\bar{t}</td>
<td>+b</td>
<td>\bar{g}</td>
<td></td>
<td>0.24</td>
</tr>
<tr>
<td>\bar{t}</td>
<td>\bar{b}</td>
<td>+g</td>
<td></td>
<td>0.06</td>
</tr>
<tr>
<td>\bar{t}</td>
<td>\bar{b}</td>
<td>\bar{g}</td>
<td></td>
<td>0.06</td>
</tr>
</tbody>
</table>
Example: Traffic

- Variables:
  - R: It rains
  - T: There is traffic
Example: Traffic

- **Variables:**
  - R: It rains
  - T: There is traffic

- **Model 1: independence**
Example: Traffic

- Variables:
  - R: It rains
  - T: There is traffic

- Model 1: independence

- Model 2: rain is conditioned on traffic
  - Why is an agent using model 2 better?
Example: Traffic

- Variables:
  - R: It rains
  - T: There is traffic

- Model 1: independence

- Model 2: rain is conditioned on traffic
  - Why is an agent using model 2 better?

- Model 3: traffic is conditioned on rain
  - Is this better than model 2?
Example: Alarm Network

- **Variables**
  - B: Burglary
  - A: Alarm goes off
  - M: Mary calls
  - J: John calls
  - E: Earthquake!
Example: Alarm Network

<table>
<thead>
<tr>
<th>B</th>
<th>P(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>+b</td>
<td>0.001</td>
</tr>
<tr>
<td>←b</td>
<td>0.999</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>E</th>
<th>P(E)</th>
</tr>
</thead>
<tbody>
<tr>
<td>+e</td>
<td>0.002</td>
</tr>
<tr>
<td>←e</td>
<td>0.998</td>
</tr>
</tbody>
</table>

| A  | J    | P(J|A) |
|----|------|------|
| +a | +j   | 0.9  |
| +a | ←j  | 0.1  |
| ←a | +j  | 0.05 |
| ←a | ←j  | 0.95 |

| A  | M    | P(M|A) |
|----|------|------|
| +a | +m   | 0.7  |
| +a | ←m  | 0.3  |
| ←a | +m  | 0.01 |
| ←a | ←m  | 0.99 |

| B  | E    | A    | P(A|B,E) |
|----|------|------|---------|
| +b | +e   | +a   | 0.95    |
| +b | +e   | ←a   | 0.05    |
| +b | ←e   | +a   | 0.94    |
| +b | ←e   | ←a   | 0.06    |
| ←b | +e   | +a   | 0.29    |
| ←b | +e   | ←a   | 0.71    |
| ←b | ←e   | +a   | 0.001   |
| ←b | ←e   | ←a   | 0.999   |
Changing Bayes’ Net Structure

- The same joint distribution can be encoded in many different Bayes’ nets

- Analysis question: given some edges, what other edges do you need to add?
  - One answer: fully connect the graph
  - Better answer: don’t make any false conditional independence assumptions
Example: Independence

- For this graph, you can fiddle with \( \setminus \) (the CPTs) all you want, but you won’t be able to represent any distribution in which the flips are dependent!

\[
\begin{align*}
X_1 & \quad X_2 \\
\text{\(P(X_1)\)} & \quad \text{\(P(X_2)\)} \\
\begin{array}{cc}
\text{h} & 0.5 \\
\text{t} & 0.5
\end{array} & \begin{array}{cc}
\text{h} & 0.5 \\
\text{t} & 0.5
\end{array}
\end{align*}
\]
Example: Coins

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

\[
P(X_1) \\
\begin{array}{c|c}
 h & 0.5 \\
 t & 0.5 \\
\end{array} \quad P(X_2) \\
\begin{array}{c|c}
 h & 0.5 \\
 t & 0.5 \\
\end{array}
\]
Example: Coins

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

\[
\begin{align*}
X_1 & \quad X_2 \\
P(X_1) & \quad P(X_2) \\
h & 0.5 & h & 0.5 \\
t & 0.5 & t & 0.5
\end{align*}
\]
Example: Coins

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

\[
P(X_1) \\
\begin{array}{cc}
h & 0.5 \\
t & 0.5 \\
\end{array}
\]

\[
P(X_2) \\
\begin{array}{cc}
h & 0.5 \\
t & 0.5 \\
\end{array}
\]

\[
P(X_1 | h) \\
\begin{array}{cc}
h | h & 0.5 \\
t | h & 0.5 \\
\end{array}
\]
Example: Coins

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

\[
\begin{align*}
P(X_1) \quad P(X_2) \\
\begin{array}{c|c}
h & 0.5 \\
t & 0.5 \\
\end{array} & \begin{array}{c|c}
h & 0.5 \\
t & 0.5 \\
\end{array}
\end{align*}
\]
Example: Coins

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

\[
P(X_1) = \begin{pmatrix}
h & 0.5 \\
t & 0.5 \end{pmatrix}
\]

\[
P(X_2) = \begin{pmatrix}
h & 0.5 \\
t & 0.5 \end{pmatrix}
\]

\[
P(X_1) \rightarrow P(X_2|X_1) = \begin{pmatrix}
h | h & 0.5 \\
t | h & 0.5 \\
h | t & 0.5 \\
t | t & 0.5
\end{pmatrix}
\]
Example: Coins

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

\[ h^{0.5} \quad t^{0.5} \]

\[ X_1 \quad X_2 \]

- Adding unneeded arcs isn’t wrong, it’s just inefficient

\[
\begin{array}{c|c}
X_1 & P(X_1) \\
\hline
h & 0.5 \\
t & 0.5 \\
\end{array}
\quad
\begin{array}{c|c}
X_2 & P(X_2) \\
\hline
h & 0.5 \\
t & 0.5 \\
\end{array}
\quad
\begin{array}{c|c|c}
X_1 & P(X_1) \\
\hline
h & h & 0.5 \\
t & t & 0.5 \\
\end{array}
\quad
\begin{array}{c|c}
X_2 | X_1 & P(X_2 | X_1) \\
\hline
h | h & 0.5 \\
t | h & 0.5 \\
h | t & 0.5 \\
t | t & 0.5 \\
\end{array}
Example: Coins

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

- Adding unneeded arcs isn’t wrong, it’s just inefficient
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.
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
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:

![Graph with nodes X, Y, and Z connected in a chain]

- Question: are X and Z independent?
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](diagram.png)

- Question: are X and Z independent?
  - Answer: no.
    - Example: low pressure causes rain, which causes traffic.
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:

Question: are X and Z independent?

- Answer: no.
  - Example: low pressure causes rain, which causes traffic.
  - Knowledge about X may change belief in Z,
  - Knowledge about Z may change belief in X (via Y)
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](image)

    - Question: are X and Z independent?
      - Answer: no.
        - Example: low pressure causes rain, which causes traffic.
        - Knowledge about X may change belief in Z,
        - Knowledge about Z may change belief in X (via Y)
        - Addendum: they *could* be independent: how?
This configuration is a “causal chain”

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

- Is X independent of Z given Y?
Causal Chains

- This configuration is a “causal chain”

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

- Is X independent of Z given Y?

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

\[ = P(z|y) \quad \text{Yes!} \]
Causal Chains

- This configuration is a “causal chain”

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

- 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
Another basic configuration: two effects of the same parent
- Are X and Z independent?

Y: Project due
X: Forum busy
Z: Lab full
Another basic configuration: two effects of the same parent
- Are X and Z independent?
- Are X and Z independent given Y?

Y: Project due
X: Forum busy
Z: Lab full
Another basic configuration: two effects of the same parent

- Are X and Z independent?
- Are 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!
Common Parent

- Another basic configuration: two effects of the same parent
  - Are X and Z independent?
  - Are X and Z independent given Y?

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

- Observing the cause blocks influence between effects.

Y: Project due
X: Forum busy
Z: Lab full

Yes!
Common Effect

- Last configuration: two causes of one effect (v-structures)
  - Are X and Z independent?

X: Raining
Z: Ballgame
Y: Traffic
Common Effect

- Last configuration: two causes of one effect (v-structures)
  - Are X and Z independent?
    - Yes: the ballgame and the rain cause traffic, but they are not correlated
    - Still need to prove they must be (try it!)

X: Raining
Z: Ballgame
Y: Traffic
Common Effect

- Last configuration: two causes of one effect (v-structures)
  - Are X and Z 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 Z independent given Y?
Common Effect

- Last configuration: two causes of one effect (v-structures)
  - Are X and Z 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 Z independent given Y?
    - 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

- Any complex example can be analyzed using these three canonical cases

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

- Solution: analyze the graph
Reachability (D-Separation)

- **Question**: Are X and Y conditionally independent given evidence vars \(\{Z\}\)?
  - Yes, if X and Y “separated” by Z
  - Look for active 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**
A much simpler D-separation test!!

- Check whether $X$ and $Y$ are disconnected in a new undirected graph $G'$ obtained from the Bayesian network using the following steps:
  - Until no nodes can be deleted do
    - Delete any leaf node $W$ from DAG $G$ as long as $W$ is not in $X$, or $Y$ or $Z$.
  - Delete all edges outgoing from nodes in $Z$
  - Remove directionality

(Not given in your book)
D-Separation

Nodes in $Z$ are shaded. Pruned nodes and edges are dotted.

Is $X = \{A, S\}$ d-separated from $Y = \{D, X\}$ by $Z = \{B, P\}$?
D-separation

Nodes in Z are shaded. Pruned nodes and edges are dotted.

Is $X = \{T, C\}$ d-separated from $Y = \{B\}$ by $Z = \{S, X\}$?
Example: Independent?

\[ R \perp B \]
\[ R \perp B | T \]
\[ R \perp B | T' \]
Example: Independent?

\[ R \perp B \quad \text{Yes} \]
\[ R \perp B | T \]
\[ R \perp B | T' \]
Example: Independent?

\[ L \perp T' | T \]
\[ L \perp B \]
\[ L \perp B | T \]
\[ L \perp B | T' \]
\[ L \perp B | T, R \]
Example: Independent?

\[ L \perp T' | T \]
\[ L \perp B \]
\[ L \perp B | T \]
\[ L \perp B | T' \]
\[ L \perp B | T, R \]
Example: Independent?

\[ L \perp T' | T \quad \text{Yes} \]
\[ L \perp B \quad \text{Yes} \]
\[ L \perp B | T \]
\[ L \perp B | T' \]
\[ L \perp B | T, R \]
Example: Independent?

\[ L \perp T' | T \quad \text{Yes} \]
\[ L \perp B \quad \text{Yes} \]
\[ L \perp B | T \]
\[ L \perp B | T' \]
\[ L \perp B | T, R \quad \text{Yes} \]
Example

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

- **Questions:**

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

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

- **Questions:**

  \[
  T \perp D
  \]

  \[
  T \perp D | R
  \]

  Yes

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