Reasoning with Bayesian Belief Networks
Overview

• Bayesian Belief Networks (BBNs) can reason with networks of propositions and associated probabilities

• BBNs encode causal associations between facts and events the propositions represent

• Useful for many AI problems
  – Diagnosis
  – Expert systems
  – Planning
  – Learning
Judea Pearl

- UCLA CS professor
- Introduced **Bayesian networks** in the 1980s
- Pioneer of probabilistic approach to AI reasoning
- First to mathematize causal modeling in empirical sciences
- Written many books on the topics, including the popular 2018 **Book of Why**
**BBN Definition**

- AKA Bayesian Network, Bayes Net
- A graphical model (as a DAG) of probabilistic relationships among a set of random variables
- Nodes are variables, links represent direct influence of one variable on another
- Nodes have prior probabilities or Conditional Probability Tables (CPTs)
Recall Bayes Rule

\[ P(H, E) = P(H \mid E)P(E) = P(E \mid H)P(H) \]

Note symmetry: can compute probability of a \emph{hypothesis given its evidence} as well as probability of \emph{evidence given hypothesis}
Simple Bayesian Network

\[ S \in \{\text{no, light, heavy}\} \quad \text{Smoking} \quad \rightarrow \quad \text{Cancer} \]

\[ C \in \{\text{none, benign, malignant}\} \]
Simple Bayesian Network

- **Smoking** variable represents person’s degree of smoking and has three possible values (no, light, heavy)
- **Cancer** variable represents person’s cancer diagnosis and has three possible values (none, benign, malignant)
Simple Bayesian Network

\[ S \in \{\text{no, light, heavy}\} \quad \text{Smoking} \quad \rightarrow \quad \text{Cancer} \quad \]

\[ C \in \{\text{none, benign, malignant}\} \]

- **tl;dr:** smoking effects cancer
- **Smoking** behavior effects the probability of cancer outcome
- **Smoking** behavior considered evidence for whether a person is likely to have cancer or not

Directed links represent "causal" relations.
Simple Bayesian Network

\[ S \in \{\text{no, light, heavy}\} \]

\( P(S=\text{no}) = 0.80 \)
\( P(S=\text{light}) = 0.15 \)
\( P(S=\text{heavy}) = 0.05 \)

\( C \in \{\text{none, benign, malignant}\} \)

Nodes without in-links have prior probabilities

Nodes with in-links have joint probability distributions

<table>
<thead>
<tr>
<th>( Smoking= )</th>
<th>no</th>
<th>light</th>
<th>heavy</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C=\text{none} )</td>
<td>0.96</td>
<td>0.88</td>
<td>0.60</td>
</tr>
<tr>
<td>( C=\text{benign} )</td>
<td>0.03</td>
<td>0.08</td>
<td>0.25</td>
</tr>
<tr>
<td>( C=\text{malignant} )</td>
<td>0.01</td>
<td>0.04</td>
<td>0.15</td>
</tr>
</tbody>
</table>
More Complex Bayesian Network

- Age
- Gender
- Exposure to Toxics
- Smoking
- Cancer
- Serum Calcium
- Lung Tumor
Nodes represent variables

• Does gender cause smoking?

• Influence might be a better term

Links represent immediate “causal” relations
More Complex Bayesian Network

- Age
- Gender
- Exposure to Toxics
- Smoking
- Cancer
- Serum Calcium
- Lung Tumor

condition
More Complex Bayesian Network

- Age
- Gender
- Exposure to Toxics
- Smoking
- Cancer
- Serum Calcium
- Lung Tumor

predispositions
More Complex Bayesian Network

- Age
- Gender
- Exposure to Toxics
- Smoking
- Cancer

observable symptoms

- Serum Calcium
- Lung Tumor
More Complex Bayesian Network

Can we predict likelihood of lung tumor given values of other 6 variables?

- Model has 7 variables
- Complete joint probability distribution will have 7 dimensions!
- Too much data required 😞
- BBN simplifies: a node has a CPT with data on itself & parents in graph
Age and Gender are independent.

\[ P(A, G) = P(G) \times P(A) \]

\[ P(A | G) = P(A) \]

\[ P(G | A) = P(G) \]

\[ P(A, G) = P(G | A) \times P(A) = P(G) \times P(A) \]

\[ P(A, G) = P(A | G) \times P(G) = P(A) \times P(G) \]
Conditional Independence

Cancer is independent of Age and Gender given Smoking

\[ P(C \mid A,G,S) = P(C \mid S) \]

If we know value of smoking, no need to know values of age or gender
Conditional Independence

Cancer is independent of Age and Gender given Smoking

• Instead of one big CPT with 4 variables, we have two smaller CPTs with 3 and 2 variables

• If all variables binary: 12 models \((2^3 + 2^2)\) rather than 16 \((2^4)\)
Conditional Independence: Naïve Bayes

Serum Calcium and Lung Tumor are dependent.

Serum Calcium is independent of Lung Tumor, given Cancer.

\[
P(L \mid SC,C) = P(L \mid C) \\
P(SC \mid L,C) = P(SC \mid C)
\]

Naïve Bayes assumption: evidence (e.g., symptoms) independent given disease; easy to combine evidence.
Explaining Away

Exposure to Toxics and Smoking are independent

Exposure to Toxics is dependent on Smoking, given Cancer

\[ P(E=\text{heavy} \mid C=\text{malignant}) > P(E=\text{heavy} \mid C=\text{malignant}, S=\text{heavy}) \]

- **Explaining away**: reasoning pattern where confirmation of one cause reduces need to invoke alternatives
- **Essence of Occam’s Razor** (prefer hypothesis with fewest assumptions)
- Relies on independence of causes
Conditional Independence

A variable (node) is conditionally independent of its non-descendants given its parents.

Cancer is independent of Age and Gender given Exposure to Toxics and Smoking.
Another non-descendant

A variable is conditionally independent of its non-descendants given its parents

Cancer is independent of Diet given Exposure to Toxics and Smoking
BBN Construction

The **knowledge acquisition** process for a BBN involves three steps

**KA1**: Choosing appropriate variables

**KA2**: Deciding on the network structure

**KA3**: Obtaining data for the conditional probability tables
KA1: Choosing variables

- Variable values: integers, reals or enumerations
- Variable should have collectively *exhaustive, mutually exclusive* values

\[ x_1 \lor x_2 \lor x_3 \lor x_4 \]
\[ \neg (x_i \land x_j) \quad i \neq j \]

- They should be values, not probabilities
Heuristic: Knowable in Principle

Example of good variables

- Weather: \{Sunny, Cloudy, Rain, Snow\}
- Gasoline: $ per gallon \{<1, 1-2, 2-3, 3-4, >4\}
- Temperature: \{\geq 100^\circ F, < 100^\circ F\}
- User needs help on Excel Charts: \{Yes, No\}
- User’s personality: \{dominant, submissive\}
KA2: Structuring

Network structure corresponding to “causality” is usually good.

Initially this uses designer’s knowledge and intuitions but can be checked with data

May be better to add suspected links than to leave out

But bigger CPT tables mean more data collection
KA3: The Numbers

- For each variable we have a table of probability of its value for values of its parents.
- For variables w/o parents, we have prior probabilities.

\[ S \in \{no, light, heavy\} \]
\[ C \in \{none, benign, malignant\} \]

<table>
<thead>
<tr>
<th>smoking priors</th>
<th>smoking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>cancer</td>
</tr>
<tr>
<td>no</td>
<td>0.96</td>
</tr>
<tr>
<td>light</td>
<td>0.03</td>
</tr>
<tr>
<td>heavy</td>
<td>0.01</td>
</tr>
</tbody>
</table>
KA3: The numbers

- Second decimal usually doesn’t matter
- Relative probabilities are important

- Zeros and ones are often enough
- Order of magnitude is typical: $10^{-9}$ vs $10^{-6}$
- Sensitivity analysis can be used to decide accuracy needed
Three kinds of reasoning

BBNs support three main kinds of reasoning:

• **Predicting** conditions given predispositions
• **Diagnosing** conditions given symptoms (and predisposing)
• **Explaining** a condition by one or more predispositions

To which we can add a fourth:

• **Deciding** on an action based on probabilities of the conditions
Predictive Inference

How likely are elderly males to get malignant cancer?

\[ P(C=\text{malignant} \mid \text{Age}>60, \text{Gender}=\text{male}) \]
How likely is an elderly male patient with high Serum Calcium to have malignant cancer?

$$P(C=\text{malignant} \mid \text{Age}>60, \text{Gender}=\text{male}, \text{Serum Calcium} = \text{high})$$
Explaining away

- If we see a lung tumor, the probability of heavy smoking and of exposure to toxics both go up.
- If we then observe heavy smoking, the probability of exposure to toxics goes back down.
Some software tools

- **Netica**: Windows app for working with Bayesian belief networks and influence diagrams
  - A commercial product, free for small networks
  - Includes graphical editor, compiler, inference engine, etc.
  - To run in OS X or Linux you need Crossover

- **Hugin**: free demo versions for Linux, Mac, and Windows are available

- Various Python packages, e.g., ...

- Aima-python code in probability4e.py
Dyspnea is difficult or labored breathing.
Same BBN model in Hugin app
Decision making

• A decision is a medical domain might be a choice of treatment (e.g., radiation or chemotherapy)
• Decisions should be made to maximize expected utility
• View decision making in terms of
  – Beliefs/Uncertainties
  – Alternatives/Decisions
  – Objectives/Utilities
Decision Problem

Should I have my party inside or outside?

Diagram:
- **in**:
  - **dry**: Regret
  - **wet**: Relieved
- **out**:
  - **dry**: Perfect!
  - **wet**: Disaster
A numerical score over all possible states allows a BBN to be used to make decisions.

<table>
<thead>
<tr>
<th>Location?</th>
<th>Weather?</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>in</td>
<td>dry</td>
<td>$50</td>
</tr>
<tr>
<td>in</td>
<td>wet</td>
<td>$60</td>
</tr>
<tr>
<td>out</td>
<td>dry</td>
<td>$100</td>
</tr>
<tr>
<td>out</td>
<td>wet</td>
<td>$0</td>
</tr>
</tbody>
</table>
Decision Making with BBNs

• Today’s weather forecast might be either sunny, cloudy or rainy
• Should you take an umbrella when you leave?
• Your decision depends only on the forecast
  – The forecast “depends on” the actual weather
• Your satisfaction depends on your decision and the weather
  – Assign a utility to each of four situations: (rain | no rain) x (umbrella, no umbrella)
Decision Making with BBNs

- Extend BBN framework to include two new kinds of nodes: **decision** and **utility**
- **Decision** node computes the expected utility of a decision given its parent(s) (e.g., forecast) and a valuation
- **Utility** node computes utility value given its parents, e.g. a decision and weather
  - Assign utility to each situations: (rain | no rain) x (umbrella, no umbrella)
  - Utility value assigned to each is probably subjective
Forecast
- Sunny: 53.5
- Cloudy: 21.5
- Rainy: 25.0

Weather
- No Rain: 70.0
- Rain: 30.0

Decide_Umbrella
- Take It: 35.0
- Leave At Home: 70.0

Satisfaction
<table>
<thead>
<tr>
<th>Weather</th>
<th>Decide_Umbrella</th>
<th>Satisfaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Rain</td>
<td>Take It</td>
<td>20</td>
</tr>
<tr>
<td>No Rain</td>
<td>Leave At Home</td>
<td>100</td>
</tr>
<tr>
<td>Rain</td>
<td>Take It</td>
<td>70</td>
</tr>
<tr>
<td>Rain</td>
<td>Leave At Home</td>
<td>0</td>
</tr>
</tbody>
</table>
Forecast

<table>
<thead>
<tr>
<th>Sunny</th>
<th>Cloudy</th>
<th>Rainy</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
</tbody>
</table>

Weather

<table>
<thead>
<tr>
<th>No Rain</th>
<th>Rain</th>
</tr>
</thead>
<tbody>
<tr>
<td>65.1</td>
<td>34.9</td>
</tr>
</tbody>
</table>

Decide_Umbrella

<table>
<thead>
<tr>
<th>Take It</th>
<th>Leave At Home</th>
</tr>
</thead>
<tbody>
<tr>
<td>37.4418</td>
<td>65.1162</td>
</tr>
</tbody>
</table>

Satisfaction
Predispositions or causes:

- **Visit To Asia**
  - Visit: 1.00
  - No visit: 99.0

- **Smoking**
  - Smoker: 50.0
  - Non smoker: 50.0

- **Tuberculosis**
  - Present: 1.04
  - Absent: 99.0

- **Lung Cancer**
  - Present: 5.50
  - Absent: 94.5

- **Bronchitis**
  - Present: 45.0
  - Absent: 55.0

- **XRay Result**
  - Abnormal: 11.0
  - Normal: 89.0

- **Dyspnea**
  - Present: 43.6
  - Absent: 56.4

**Chest Clinic**
Distributed by Norsys Software Corp.
Conditions or diseases

Visits to Asia
- Visit
  - Present: 1.04
  - Absent: 99.0
- No Visit
  - Present: 99.0
  - Absent: 1.00

Smoking
- Present: 5.50
  - Absent: 94.5

Tuberculosis
- Present: 45.0
  - Absent: 55.0

Lung Cancer
- Present: 5.50
  - Absent: 94.5

Bronchitis
- Present: 45.0
  - Absent: 55.0

Tuberculosis or Cancer
- True: 6.48
  - False: 93.5

XRay Result
- Abnormal: 11.0
  - Normal: 89.0

Dyspnea
- Present: 43.6
  - Absent: 56.4
Chest Clinic

Functional Node

Tuberculosis or Cancer
- true: 6.48%
- false: 93.5%

XRay Result
- abnormal: 11.0%
- normal: 89.0%

Dyspnea
- present: 43.6%
- absent: 56.4%

Smoking
- smoker: 50.0%
- non smoker: 50.0%

Visit To Asia
- visit: 1.00%
- no visit: 99.0%
Symptoms or effects

Dyspnea is shortness of breath