Reasoning with Bayesian Belief Networks
Overview

- Bayesian Belief Networks (BBNs) can reason with networks of propositions and associated probabilities
- Useful for many AI problems
  - Diagnosis
  - Expert systems
  - Planning
  - Learning
BBN Definition

• AKA Bayesian Network, Bayes Net
• A graphical model (as a DAG) of probabilistic relationships among a set of random variables
• Links represent direct influence of one variable on another

![Diagram of Bayesian Network][1]

[1]: source
Recall Bayes Rule

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

\[ P(H \mid E) = \frac{P(E \mid H)P(H)}{P(E)} \]

Note symmetry: can compute probability of a hypothesis given its evidence as well as probability of 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

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

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

Nodes represent variables

Links represent "causal" relations
**Simple Bayesian Network**

$S \in \{\text{no, light, heavy}\}$  

Prior probability of $S$

<table>
<thead>
<tr>
<th>$P(S=\text{no})$</th>
<th>0.80</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(S=\text{light})$</td>
<td>0.15</td>
</tr>
<tr>
<td>$P(S=\text{heavy})$</td>
<td>0.05</td>
</tr>
</tbody>
</table>

$C \in \{\text{none, benign, malignant}\}$

Joint distribution of $S$ and $C$

<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>

Nodes with no in-links have prior probabilities

Nodes with in-links have joint probability distributions
Simple Bayesian Network

Nodes represent variables.

Prior probability of $S$

<table>
<thead>
<tr>
<th>$P(S=\text{no})$</th>
<th>0.80</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(S=\text{light})$</td>
<td>0.15</td>
</tr>
<tr>
<td>$P(S=\text{heavy})$</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Joint distribution of $S$ and $C$

<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
More Complex Bayesian Network

Nodes represent variables

- Does gender cause smoking?
- Influence might be a better term

Links represent “causal” relations
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

Serum Calcium

Lung Tumor

condition
More Complex Bayesian Network

Age → Gender

Exposure to Toxics → Smoking

Cancer

Serum Calcium
Lung Tumor

observable symptoms
More Complex Bayesian Network

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

- Model has 7 variables
- Complete joint probability distribution will have 7 dimensions!
- A lot of data must be collected
Independence

Age and Gender are independent.

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

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

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

There is no path between them in the graph.
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, we don’t need to know values of age or gender
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: Cents per gallon {0,1,2...}
- Temperature: {≥ 100° F, < 100° 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 the designer’s knowledge but can be checked with data.
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

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

<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></td>
</tr>
<tr>
<td>light</td>
<td></td>
</tr>
<tr>
<td>heavy</td>
<td></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
How likely are elderly males to get malignant cancer?

$$P(C=\text{malignant} \mid \text{Age}>60, \text{Gender}=\text{male})$$
Predictive and diagnostic combined

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
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?

- **in**
  - dry: Regret
  - wet: Relieved

- **out**
  - dry: Perfect!
  - wet: Disaster
Value Function

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>

Using $ for the value helps our intuition.
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.

• **Hugin**: free demo version for linux, mac, windows is available
Same BBN model in Hugin app
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

- 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
Distributed by Norsys Software Corp
Conditions or diseases

Visit To Asia
- visit: 1.00
- no visit: 99.0

Smoking

Tuberculosis
- present: 1.04
- absent: 99.0

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
Distributed by Norsys Software Corp
Functional Node
Symptoms or effects

Dyspnea is shortness of breath
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
### Satisfaction Table (in net N3_Umbrella)

**Node:** 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:
- Sunny: 0
- Cloudy: 100
- Rainy: 0

Weather:
- No Rain: 65.1
- Rain: 34.9

Decide_Umbrella:
- Take It: 37.4418
- Leave At Home: 65.1162

Satisfaction