Administrivia

- Project 2 due Sun Mar 25 11:59PM
- ULW second draft due April 1
- Posters April 24 & 26
Bayesian Networks
(Probabilistic Inference)
Sources of Uncertainty

Partial Observability and Nondeterminism
Probability

Probability: Degree of belief in possible world

\[ 0 \leq P(\omega) \leq 1 \]

\[ \sum_{\omega \in \Omega} P(\omega) = 1 \]

Probability of a proposition: Conditional probability:

\[ P(\phi) = \sum_{\omega \in \phi} P(\omega) \]

\[ P(a \mid b) = \frac{P(a \land b)}{P(b)} \text{ when } P(b) > 0 \]

\[ P(a \land b) = P(a \mid b)P(b) \]
The Language of Probability Assertions

- Random variables (and domains)
- Atomic propositions:
  - $Var = value$ (or $<, \leq, >, \geq$)
- Connectives
Probability Distributions

• Describe the probabilities of all the possible values of a random variable

\[ P(\text{Weather} = \text{sunny}) = 0.6 \]
\[ P(\text{Weather} = \text{rain}) = 0.1 \]
\[ P(\text{Weather} = \text{cloudy}) = 0.29 \]
\[ P(\text{Weather} = \text{snow}) = 0.01 \]

\[ \mathbf{P}(\text{Weather}) = \langle 0.6, 0.1, 0.29, 0.01 \rangle \]
**Full Joint Probability Distribution**

\[ P(\text{Cavity, Toothache, Weather}) \]

<table>
<thead>
<tr>
<th></th>
<th>toothache</th>
<th>¬toothache</th>
</tr>
</thead>
<tbody>
<tr>
<td>cavity</td>
<td>sunny</td>
<td></td>
</tr>
<tr>
<td>¬cavity</td>
<td>rain</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cloudy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>snow</td>
<td></td>
</tr>
</tbody>
</table>
Programming Language for Knowledge 3.0

- Probabilities
- Language of probability assertions
- Distributions, conditional distributions, joint distributions, full joint distributions
Probabilistic Inference
Probabilistic Inference

- Computing posterior probabilities for propositions given prior probabilities and observed evidence

- or: Given priors and evidence, compute probabilities given evidence (posteriors)
\[ P(\text{Cavity}, \text{Toothache}, \text{Catch}) \]

<table>
<thead>
<tr>
<th></th>
<th>toothache</th>
<th>(\neg\text{toothache})</th>
</tr>
</thead>
<tbody>
<tr>
<td>catch</td>
<td>0.108</td>
<td>0.012</td>
</tr>
<tr>
<td>(\neg\text{catch})</td>
<td>0.072</td>
<td>0.008</td>
</tr>
<tr>
<td>(\neg\text{cavity})</td>
<td>0.016</td>
<td>0.064</td>
</tr>
<tr>
<td>(\text{cavity})</td>
<td>0.144</td>
<td>0.576</td>
</tr>
</tbody>
</table>
Probability of a Proposition

\[ P(\text{cavity} \lor \text{toothache}) = 0.28 \]

<table>
<thead>
<tr>
<th></th>
<th>toothache</th>
<th>(\neg\text{toothache})</th>
</tr>
</thead>
<tbody>
<tr>
<td>\text{catch}</td>
<td>0.108</td>
<td>0.012</td>
</tr>
<tr>
<td>(\neg\text{catch})</td>
<td>0.072</td>
<td>0.008</td>
</tr>
<tr>
<td>\text{cavity}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>\text{catch}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.016</td>
<td>0.064</td>
<td></td>
</tr>
<tr>
<td>(\neg\text{catch})</td>
<td>0.144</td>
<td>0.576</td>
</tr>
</tbody>
</table>
Probability of a Proposition

\[ P(\text{cavity} \land \text{toothache}) = 0.12 \]

<table>
<thead>
<tr>
<th></th>
<th>toothache</th>
<th>(\neg\text{toothache})</th>
</tr>
</thead>
<tbody>
<tr>
<td>catch</td>
<td>0.108</td>
<td>0.012</td>
</tr>
<tr>
<td>(\neg\text{cavity})</td>
<td>0.016</td>
<td>0.064</td>
</tr>
</tbody>
</table>
Conditional Probability

\[
P(cavity \mid toothache) = \frac{P(cavity \land toothache)}{P(toothache)}
\]

<table>
<thead>
<tr>
<th></th>
<th>toothache</th>
<th>\neg toothache</th>
</tr>
</thead>
<tbody>
<tr>
<td>catch</td>
<td>0.108</td>
<td>0.012</td>
</tr>
<tr>
<td>\neg cavity</td>
<td>0.016</td>
<td>0.064</td>
</tr>
<tr>
<td>cavity</td>
<td>0.072</td>
<td>0.008</td>
</tr>
<tr>
<td>\neg catch</td>
<td>0.144</td>
<td>0.576</td>
</tr>
</tbody>
</table>
Conditional Probability

\[ P(\text{cavity} \mid \text{toothache}) = \frac{P(\text{cavity} \land \text{toothache})}{P(\text{toothache})} = 0.12 \]

<table>
<thead>
<tr>
<th></th>
<th>toothache</th>
<th>¬toothache</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Catch</td>
<td>0.108</td>
<td>0.012</td>
<td>0.072</td>
<td>0.008</td>
<td></td>
<td></td>
</tr>
<tr>
<td>¬Cavity</td>
<td>0.016</td>
<td>0.064</td>
<td>0.144</td>
<td>0.576</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Conditional Probability

\[ P(\text{cavity} \mid \text{toothache}) = \frac{P(\text{cavity} \land \text{toothache})}{P(\text{toothache})} = \frac{0.12}{0.2} \]

<table>
<thead>
<tr>
<th></th>
<th>toothache</th>
<th>(\neg\text{toothache})</th>
</tr>
</thead>
<tbody>
<tr>
<td>catch</td>
<td>0.108</td>
<td>0.072</td>
</tr>
<tr>
<td>(\neg\text{catch})</td>
<td>0.012</td>
<td>0.008</td>
</tr>
<tr>
<td>cavity</td>
<td>0.016</td>
<td>0.144</td>
</tr>
<tr>
<td>(\neg\text{cavity})</td>
<td>0.064</td>
<td>0.576</td>
</tr>
</tbody>
</table>
Conditional Probability

\[ P(\text{cavity} \mid \text{toothache}) = \frac{P(\text{cavity} \land \text{toothache})}{P(\text{toothache})} = 0.6 \]

<table>
<thead>
<tr>
<th></th>
<th>toothache</th>
<th>¬toothache</th>
</tr>
</thead>
<tbody>
<tr>
<td>catch</td>
<td>0.108</td>
<td>0.012</td>
</tr>
<tr>
<td>¬catch</td>
<td>0.016</td>
<td>0.064</td>
</tr>
<tr>
<td>cavity</td>
<td>0.072</td>
<td>0.008</td>
</tr>
<tr>
<td>¬cavity</td>
<td>0.144</td>
<td>0.576</td>
</tr>
</tbody>
</table>
Conditional Probability

\[ P(\text{cavity} \mid \text{toothache}) = \frac{P(\text{cavity} \land \text{toothache})}{P(\text{toothache})} = 0.6 \]

\[ P(\neg \text{cavity} \mid \text{toothache}) = \frac{P(\neg \text{cavity} \land \text{toothache})}{P(\text{toothache})} \]
Conditional Probability

\[
P(\neg \text{cavity} \mid \text{toothache}) = \frac{P(\neg \text{cavity} \land \text{toothache})}{P(\text{toothache})} = 0.4
\]

<table>
<thead>
<tr>
<th></th>
<th>toothache</th>
<th>(\neg)toothache</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>catch</td>
<td>(\neg)catch</td>
</tr>
<tr>
<td>cavity</td>
<td>0.108</td>
<td>0.012</td>
</tr>
<tr>
<td>(\neg)cavity</td>
<td>0.016</td>
<td>0.064</td>
</tr>
</tbody>
</table>
Conditional Probability

\[
P(\text{cavity} \mid \text{toothache}) = \frac{P(\text{cavity} \land \text{toothache})}{P(\text{toothache})} = 0.6
\]

\[
P(\neg \text{cavity} \mid \text{toothache}) = \frac{P(\neg \text{cavity} \land \text{toothache})}{P(\text{toothache})} = 0.4
\]

\[
\mathbf{P}(\text{Cavity} \mid \text{toothache}) = \langle 0.6, 0.4 \rangle
\]
Probabilistic Inference

\[ P(\text{cavity} \mid \text{toothache}) = \frac{P(\text{cavity} \land \text{toothache})}{P(\text{toothache})} = 0.6 \]

\[ P(\neg \text{cavity} \mid \text{toothache}) = \frac{P(\neg \text{cavity} \land \text{toothache})}{P(\text{toothache})} = 0.4 \]
Normalization

\[ P(\text{cavity} \mid \text{toothache}) = \alpha \ P(\text{cavity} \land \text{toothache}) = \alpha \ 0.12 \]

\[ P(\neg \text{cavity} \mid \text{toothache}) = \alpha \ P(\neg \text{cavity} \land \text{toothache}) = \alpha \ 0.08 \]
Normalization

\[ P(\text{cavity} \mid \text{toothache}) = \alpha \ P(\text{cavity} \land \text{toothache}) = \alpha \ 0.12 \]

\[ P(\neg \text{cavity} \mid \text{toothache}) = \alpha \ P(\neg \text{cavity} \land \text{toothache}) = \alpha \ 0.08 \]

\[ \mathbf{P}(\text{Cavity} \mid \text{toothache}) = \alpha \langle 0.12, 0.08 \rangle \]
Normalization

\[ P(\text{cavity} \mid \text{toothache}) = \alpha \, P(\text{cavity} \land \text{toothache}) = \alpha \, 0.12 \]

\[ P(\neg \text{cavity} \mid \text{toothache}) = \alpha \, P(\neg \text{cavity} \land \text{toothache}) = \alpha \, 0.08 \]

\[ P(\text{Cavity} \mid \text{toothache}) = \frac{1}{0.12 + 0.08} \langle 0.12, 0.08 \rangle \]
Normalization

\[ P(\text{cavity} \mid \text{toothache}) = \alpha \ P(\text{cavity} \land \text{toothache}) = \alpha \ 0.12 \]

\[ P(\neg\text{cavity} \mid \text{toothache}) = \alpha \ P(\neg\text{cavity} \land \text{toothache}) = \alpha \ 0.08 \]

\[ P(\text{Cavity} \mid \text{toothache}) = \langle 0.6, 0.4 \rangle \]

We didn’t know \( P(\text{toothache}) \)!
Inference
(Single Variable)

\[ P(X \mid e) = \alpha \begin{align*} P(X, e) &= \alpha \sum_y P(X, e, y) \end{align*} \]

Query variable \( X : Domain(X) = \{x_1, \ldots, x_m\} \)

Evidence variables \( E : \{E_1, \ldots, E_k\} \)

Observations \( e : \{e_1, \ldots, e_k\} \) s.t. \( E_i = e_i \)

Unobserved variables \( Y : \{Y_1, \ldots, Y_l\} \)

\[ Domain(Y_i) = \{y_{i,1}, \ldots, y_{i,n_i}\} \]
Inference (Single Variable)

\[ P(X \mid e) = \alpha P(X, e) = \alpha \sum_y P(X, e, y) \]

For each possible value \( x_i \) for \( X \)

For each possible combination of values \( y \) for \( Y \)

Add \( P(x_i, e, y) \)

Result: vector \( P(X \mid e) = \langle P(x_i \mid e) \rangle = \langle P(X = x_i \mid e) \rangle \)
Inference
(Single Variable)

\[ P(X \mid e) = \alpha P(X, e) = \alpha \sum_y P(X, e, y) \]

For each possible value \( x_i \) for \( X \)

For each possible combination of values \( y \) for \( Y \)

Add \( P(x_i, e, y) \)

\[ P(X = x_i, E_1 = e_1, \ldots, E_k = e_k, \ldots, Y_1 = y_{1, i_1}, \ldots, Y_l = y_{l, i_l}) \]

Result: vector \( P(X \mid e) = \langle P(x_i \mid e) \rangle = \langle P(X = x_i \mid e) \rangle \)
Inference
(Single Variable)

\[ P(X \mid e) = \alpha \, P(X, e) = \alpha \sum_y P(X, e, y) \]

```c
int m;   // Number of values in domain of query variable X
int k;   // Number of evidence variables E
int[k] e; // e[i]: index of i’th evidence value in domain of Ei
int l;   // Number of unobserved variables Y
int[l] n; // n[i]: number of values in domain of Yi
double[l][n[l]] D; // D[i][j]: j’th value of domain for Yi

for i from 1 to m
    PXe[i] = 0
    for i1 from 1 to n[1]
        for i2 from 1 to n[2]
            ...          ...
            for il from 1 to n[l]
                PXe[i] += FJPD[i][e[1]]...[e[k]][D[1][i1]]...[D[l][il]]
    return PXe // vector of length m
```

Full joint prob. dist.

\( l \) nested loops
Inference (Single Variable)

\[ P(X \mid e) = \alpha \quad P(X, e) = \alpha \sum_y P(X, e, y) \]

Time Complexity: \( O(m^n) \)

Space Complexity: \( O(m^n) \)
Bayesian Networks
(Probabilistic Inference)
Conditional Probability

\[ P(a \mid b) = \frac{P(a \land b)}{P(b)} \quad \text{when } P(b) > 0 \]
Product Rule

\[ P(a \land b) = P(a \mid b)P(b) \quad P(b \land a) = P(b \mid a)P(a) \]

\[ P(a \mid b)P(b) = P(b \mid a)P(a) \]
Bayes’ Rule

\[ P(b \mid a) = \frac{P(a \mid b)P(b)}{P(a)} \]

Thomas Bayes
(c. 1702 – 1761)
Bayes’ Rule

\[ P(Y \mid X) = \frac{P(X \mid Y)P(Y)}{P(X)} \]

\[ P(Y \mid X) = \alpha \ P(X \mid Y)P(Y) \]
Causal and Diagnostic Knowledge

Causal knowledge: $P(\text{effect} \mid \text{cause})$

Diagnostic knowledge: $P(\text{cause} \mid \text{effect})$
Causal and Diagnostic Knowledge

Causal knowledge: $P(\text{symptom} \mid \text{disease})$

Diagnostic knowledge: $P(\text{disease} \mid \text{symptom})$
Bayesian Diagnosis

\[ P(\text{disease} \mid \text{symptom}) = \frac{P(\text{symptom} \mid \text{disease})P(\text{disease})}{P(\text{symptom})} \]
Meningitis causes a stiff neck 70% of the time

\[ P(\text{stiffneck} \mid \text{meningitis}) = 0.7 \]

Prior probability of meningitis

\[ P(\text{meningitis}) = 0.00002 \]

Prior probability of stiff neck

\[ P(\text{stiffneck}) = 0.01 \]

\[
P(\text{meningitis} \mid \text{stiffneck}) = \frac{P(\text{stiffneck} \mid \text{meningitis})P(\text{meningitis})}{P(\text{stiffneck})}\
\]

\[ = \frac{0.7 \times 0.00002}{0.01} \]

\[ = 0.0014 \]
\textit{toothache} \quad (\text{Toothache} = \text{True})

\textit{catch} \quad (\text{Catch} = \text{True})

P(\text{Cavity} \mid \text{toothache} \land \text{catch})
Combining Evidence

\[ P(Cavity \mid toothache \land catch) \]
Combining Evidence

\[ P(Cavity \mid toothache \land catch) \]

\[ = \alpha \langle 0.180, 0.016 \rangle \approx \langle 0.871, 0.129 \rangle \]

<table>
<thead>
<tr>
<th></th>
<th>toothache</th>
<th>\neg toothache</th>
</tr>
</thead>
<tbody>
<tr>
<td>catch</td>
<td>0.108</td>
<td>0.012</td>
</tr>
<tr>
<td>\neg catch</td>
<td>0.072</td>
<td>0.008</td>
</tr>
<tr>
<td>cavity</td>
<td>0.016</td>
<td>0.064</td>
</tr>
<tr>
<td>\neg cavity</td>
<td>0.144</td>
<td>0.576</td>
</tr>
</tbody>
</table>
Combining Evidence

\[ P(Cavity \mid toothache \land catch) \]
\[ = P(toothache \land catch \mid Cavity)P(Cavity) \]
Combining Evidence

• In general, if there are $n$ possible evidence variables, then there are $O(2^n)$ possible combinations of observed values for which we would need to know the conditional probabilities.
Independence

Cavity

Weather
Independence

\[ P(\text{cavity} \mid \text{sunny}) = P(\text{cavity}) \]

\[ P(\text{rain} \mid \text{cavity}) = P(\text{rain}) \]

\[ P(\text{Cavity} \mid \text{Weather}) = P(\text{Cavity}) \]

\[ P(\text{Weather} \mid \text{Cavity}) = P(\text{Weather}) \]
Independence

\[ P(a \mid b) = P(a) \]
\[ P(b \mid a) = P(b) \]
\[ P(a \land b) = P(a)P(b) \]

\[ P(a \mid b) = \frac{P(a \land b)}{P(b)} = P(a) \]
\[ P(b \mid a) = \frac{P(b \land a)}{P(a)} = P(b) \]
Independence

\[ P(X \mid Y) = P(X) \]
\[ P(Y \mid X) = P(Y) \]
\[ P(X, Y) = P(X)P(Y) \]
Independence

\[ \mathbf{P}(X) = \langle x_1, x_2, \ldots, x_n \rangle \]
\[ \mathbf{P}(Y) = \langle y_1, y_2, \ldots, y_m \rangle \]

<table>
<thead>
<tr>
<th></th>
<th>( x_1 )</th>
<th>( x_2 )</th>
<th>( \ldots )</th>
<th>( x_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( y_1 )</td>
<td>( x_1 y_1 )</td>
<td>( x_2 y_1 )</td>
<td>( \ldots )</td>
<td>( x_n y_1 )</td>
</tr>
<tr>
<td>( y_2 )</td>
<td>( x_1 y_2 )</td>
<td>( x_2 y_2 )</td>
<td>( \ldots )</td>
<td>( x_n y_2 )</td>
</tr>
<tr>
<td>( \ldots )</td>
<td>( \ldots )</td>
<td>( \ldots )</td>
<td>( \ldots )</td>
<td>( \ldots )</td>
</tr>
<tr>
<td>( y_m )</td>
<td>( x_1 y_m )</td>
<td>( x_2 y_m )</td>
<td>( \ldots )</td>
<td>( x_n y_m )</td>
</tr>
</tbody>
</table>
Independence

\[ P(X \mid Y) = P(X) \]

\[ P(Y \mid X) = P(Y) \]

\[ P(X, Y) = P(X)P(Y) \]

Can compute \( n \times m \) probabilities from \( n + m \) probabilities (if random variables are independent)
toothache
toothache (Toothache = True)

catch
catch (Catch = True)

\[ P(Cavity \mid \text{toothache} \land \text{catch}) \]
Conditional Independence

- Both *toothache* and *catch* are caused by a cavity, but neither has a direct effect on the other.

- The variables are independent given the presence or absence of a cavity.
Conditional Independence

\[ P(\text{toothache} \land \text{catch} \mid \text{Cavity}) = P(\text{toothache} \mid \text{Cavity})P(\text{catch} \mid \text{Cavity}) \]
Conditional Independence

\[ P(\text{Toothache, Catch} \mid \text{Cavity}) = P(\text{Toothache} \mid \text{Cavity})P(\text{Catch} \mid \text{Cavity}) \]
Conditional Independence

\[
P(X \mid Y, Z) = P(X \mid Z) \\
P(Y \mid X, Z) = P(Y \mid Z) \\
P(X, Y \mid Z) = P(X \mid Z)P(Y \mid Z)
\]
Combining Evidence

\[ P(Cavity \mid toothache \land catch) = \]
\[ P(toothache \land catch \mid Cavity) = P(toothache \mid Cavity)P(catch \mid Cavity) \]
Bayes’ Rule

\[ P(Y \mid X) = \frac{P(X \mid Y)P(Y)}{P(X)} \]

\[ P(Y \mid X) = \alpha \ P(X \mid Y)P(Y) \]
Combining Evidence

\[ P(Cavity \mid \text{toothache} \land \text{catch}) \]

\[ P(\text{toothache} \land \text{catch} \mid Cavity) = P(\text{toothache} \mid Cavity)P(\text{catch} \mid Cavity) \]
Combining Evidence

\[ P(Cavity \mid toothache \land catch) = \alpha P(toothache \mid Cavity)P(catch \mid Cavity)P(Cavity) \]
Combining Evidence

• For $n$ symptoms (e.g., Toothache, Catch) that are all conditionally independent given a disease (e.g., Cavity), we need $O(n)$ probabilities rather than $O(2^n)$

• Representation scales to larger problems

• More likely to be available than absolute independence assumptions
Example Section 13.6
You can do it!
Probabilistic Reasoning

- Full joint distribution: intractable as problem grows
- Independence assumptions reduce number of probabilities required to represent full joint distribution
- Next: Develop a data structure that represents the dependencies among random variables and can be used to compute any full joint distribution
Bayesian Networks

Random Variables

Cavity

Toothache

Catch

“has direct influence on”

conditionally independent given parents
Bayesian Networks

Weather → Cavity → Toothache → Catch
Bayesian Networks

\[
P(\text{Toothache} | \text{Cavity})
\]
\[
P(\text{Catch} | \text{Cavity})
\]
\[
P(\text{Cavity})
\]

Conditional Probability Distributions
Bayesian Networks

Prior Probability Distribution

\[
P(\text{Toothache} | \text{Cavity})
\]

\[
P(\text{Catch} \mid \text{Cavity})
\]

\[
P(\text{Cavity})
\]
Bayesian Networks

- Each node corresponds to a random variable
- There is a link from $X$ to $Y$ if $X$ has a direct influence on $Y$ (no cycles $\Rightarrow$ DAG)
- The node for $X_i$ stores the conditional distribution $P(X_i \mid \text{Parents}(X_i))$
Bayesian Networks
How-To

• Select random variables required to model the domain

• Add links from causes to effects ("directly influences")

• No cycles (see book)

• Write down conditional probability distributions for $P(X_i \mid Parents(X_i))$
Representing Conditional Distributions

- Burglary: $P(B) = .001$
- Earthquake: $P(E) = .002$
- Alarm:
  - $P(A|t,t) = .95$
  - $P(A|t,f) = .94$
  - $P(A|f,t) = .29$
  - $P(A|f,f) = .001$
- JohnCalls:
  - $P(J|t) = .90$
  - $P(J|f) = .05$
- MaryCalls:
  - $P(M|t) = .70$
  - $P(M|f) = .01$
SO WHAT!
THE GOOD, THE MAD AND THE UGLY
EDITED BY STEFANER CHIBAJI
Inference in Bayesian Networks

- Query variable $X$
- Evidence variables $E_1, \ldots, E_m$
- Observed values: $e = \langle e_1, \ldots, e_m \rangle$
- Non-evidence, non-query (“hidden”) variables: $Y$
- Compute: $P(X | e)$
Inference in Bayesian Networks

\[ P(X \mid e) = \alpha P(X, e) = \alpha \sum_y P(X, e, y) \]

Entry in full joint distribution

\[ P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \mid \text{parents}(X_i)) \]
Inference in Bayesian Networks

- A query can be answered from a Bayesian Network by computing sums of products of conditional probabilities from the network.
A diagram showing the relationships between events:

- **Burglary** with a probability of 0.001.
- **Earthquake** with a probability of 0.002.

The event **Alarm** is influenced by

- **Burglary** and **Earthquake**.

The outcomes of **Alarm** depend on

- Two conditions: **B** and **E**.

<table>
<thead>
<tr>
<th>B</th>
<th>E</th>
<th>P(A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>t</td>
<td>0.95</td>
</tr>
<tr>
<td>t</td>
<td>f</td>
<td>0.94</td>
</tr>
<tr>
<td>f</td>
<td>t</td>
<td>0.29</td>
</tr>
<tr>
<td>f</td>
<td>f</td>
<td>0.001</td>
</tr>
</tbody>
</table>

The events **JohnCalls** and **MaryCalls** are influenced by

- **Alarm**.

The outcomes of **JohnCalls** depend on

<table>
<thead>
<tr>
<th>A</th>
<th>P(J)</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>0.90</td>
</tr>
<tr>
<td>f</td>
<td>0.05</td>
</tr>
</tbody>
</table>

The outcomes of **MaryCalls** depend on

<table>
<thead>
<tr>
<th>A</th>
<th>P(M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>0.70</td>
</tr>
<tr>
<td>f</td>
<td>0.01</td>
</tr>
<tr>
<td>B</td>
<td>E</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>t</td>
<td>t</td>
</tr>
<tr>
<td>t</td>
<td>f</td>
</tr>
<tr>
<td>f</td>
<td>t</td>
</tr>
<tr>
<td>f</td>
<td>f</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A</th>
<th>P(J)</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>.90</td>
</tr>
<tr>
<td>f</td>
<td>.05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A</th>
<th>P(M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>.70</td>
</tr>
<tr>
<td>f</td>
<td>.01</td>
</tr>
</tbody>
</table>
$P(X_i \mid Parents(X_i))$
### Events

- **Burglary**
  - $P(B) = 0.001$

- **Earthquake**
  - $P(E) = 0.002$

- **Alarm**

- **JohnCalls**
  - $A$ $P(J)$
    - $t$ 0.90
    - $f$ 0.05

- **MaryCalls**
  - $A$ $P(M)$
    - $t$ 0.70
    - $f$ 0.01

### Conditional Probabilities

<table>
<thead>
<tr>
<th>$B$</th>
<th>$E$</th>
<th>$P(A)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t$</td>
<td>$t$</td>
<td>0.95</td>
</tr>
<tr>
<td>$t$</td>
<td>$f$</td>
<td>0.94</td>
</tr>
<tr>
<td>$f$</td>
<td>$t$</td>
<td>0.29</td>
</tr>
<tr>
<td>$f$</td>
<td>$f$</td>
<td>0.001</td>
</tr>
</tbody>
</table>
\[ P(\text{Burglary} \mid \text{JohnCalls} = \text{True}, \text{MaryCalls} = \text{True}) \]

\[ P(B \mid j, m) \]
\[ P(B \mid j, m) = \alpha P(B, j, m) = \alpha \sum_{e} \sum_{a} P(B, j, m, e, a) \]
\[ P(B \mid j, m) = \alpha P(B, j, m) = \alpha \sum_e \sum_a P(B, j, m, e, a) \]

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

\[ P(b \mid j, m) = \alpha \sum_e \sum_a P(b)P(e)P(a \mid b, e)P(j \mid a)P(m \mid a) \]
\[ P(b \mid j, m) = \alpha \sum_e \sum_a P(b) P(e) P(a \mid b, e) P(j \mid a) P(m \mid a) \]
\[ P(B \mid j, m) = \alpha P(B, j, m) = \alpha \sum_{e} \sum_{a} P(B, j, m, e, a) \]

\[ P(b \mid j, m) = \alpha \sum_{e} \sum_{a} P(b)P(e)P(a \mid b, e)P(j \mid a)P(m \mid a) \]

\[ O(n2^n) \]
\[ P(B \mid j, m) = \alpha \prod_{e, a} P(B, j, m) = \alpha \sum_{e} \sum_{a} P(B, j, m, e, a) \]

\[ P(b \mid j, m) = \alpha \sum_{e} \sum_{a} P(b)P(e)P(a \mid b, e)P(j \mid a)P(m \mid a) \]

\[ P(b \mid j, m) = \alpha P(b) \sum_{e} P(e) \sum_{a} P(a \mid b, e)P(j \mid a)P(m \mid a) \]
\[ \mathbf{P}(B \mid j, m) = \alpha \langle 0.00059224, 0.0014919 \rangle \approx \langle 0.284, 0.716 \rangle \]
Exact Inference in BNs

Time Complexity: $O(2^n)$

Space Complexity: $O(n)$
Exact Inference in BNs

- Exact inference in BNs is NP-hard
- Can be shown to be as hard as computing the number of satisfying assignments of a propositional logic formula => \#P-hard
Bayesian Networks

Summary

• Independence assumptions make probabilistic inference easier

• By factoring the joint distribution

• Bayesian Networks encode independence assumptions among random variables

• And store conditional probabilities

• Represent your problem as BN: WRITE NO CODE!
Bayesian Networks

Summary

• Exact inference with BNs is still hard

• But we can do approximate inference efficiently (next time)

• We can learn the conditional probabilities required to do inference from data (in a few weeks)
For Next Time:
14.4-14.5; 14.7 fyi