Probabilistic Graphical Models

Representation
Bayesian Networks

Semantics & Factorization
- Grade
- Course Difficulty
- Student Intelligence
- Student SAT
- Reference Letter

\[ P(G,D,I,S,L) \]
Chain Rule for Bayesian Networks

\[
P(D, I, G, S, L) = P(D) P(I) P(G|I,D) P(S|I) P(L|G)
\]

Distribution defined as a product of factors!
P(d₀, i¹, g³, s¹, l¹) = 0.6 * 0.3 * 0.2 * 0.95 * 0.8
Bayesian Network

• A Bayesian network is:
  – A directed acyclic graph (DAG) $G$ whose nodes represent the random variables $X_1, \ldots, X_n$
  – For each node $X_i$ a CPD $P(X_i \mid \text{Par}_G(X_i))$

• The BN represents a joint distribution via the chain rule for Bayesian networks

$$P(X_1, \ldots, X_n) = \prod_i P(X_i \mid \text{Par}_G(X_i))$$
BN Is a Legal Distribution: $P \geq 0$

- $P$ is a product of CPDs.
- CPDs are non-negative.
BN Is a Legal Distribution: $\sum P = 1$

\[
\sum_{D,I,G,S,L} P(D,I,G,S,L) = \sum_{D,I,G,S,L} P(D) \cdot P(I) \cdot P(G|I,D) \cdot P(S|I) \cdot P(L|G)
\]

chain rule

\[
= \sum_{D,I,G,S} P(D) \cdot P(I) \cdot P(G|I,D) \cdot P(S|I) \cdot \sum_{L} P(L|G)
\]

\[
= \sum_{D,I,G,S} P(D) \cdot P(I) \cdot P(G|I,D) \cdot P(S|I)
\]

\[
= \sum_{D,I,G} P(D) \cdot P(I) \cdot P(G|I,D) \cdot \sum_{S} P(S|I)
\]

\[
= \sum_{D,I} P(D) \cdot P(I) \cdot \sum_{G} P(G|I,D)
\]
P Factorizes over G

Let $G$ be a graph over $X_1, \ldots, X_n$.
P factorizes over $G$ if

$$P(X_1, \ldots, X_n) = \prod_i P(X_i | \text{Par}_G(X_i))$$
Genetic Inheritance

Chromosome:

- AA, AB, AO, BO, BB, OO
- A, B, AB, O

Family Tree:

- Homer
  - Bart
  - Lisa
  - Maggie
- Marge
- Clancy
- Jackie
- Selma

- Genotype
- Phenotype
BNs for Genetic Inheritance
The Student Network

<table>
<thead>
<tr>
<th>d^0</th>
<th>d^1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6</td>
<td>0.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>i^0</th>
<th>i^1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.7</td>
<td>0.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>g^1</th>
<th>g^2</th>
<th>g^3</th>
</tr>
</thead>
<tbody>
<tr>
<td>i^0, d^0</td>
<td>0.3</td>
<td>0.4</td>
</tr>
<tr>
<td>i^0, d^1</td>
<td>0.05</td>
<td>0.25</td>
</tr>
<tr>
<td>i^1, d^0</td>
<td>0.9</td>
<td>0.08</td>
</tr>
<tr>
<td>i^1, d^1</td>
<td>0.5</td>
<td>0.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>s^0</th>
<th>s^1</th>
</tr>
</thead>
<tbody>
<tr>
<td>i^0</td>
<td>0.95</td>
</tr>
<tr>
<td>i^1</td>
<td>0.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>l^0</th>
<th>l^1</th>
</tr>
</thead>
<tbody>
<tr>
<td>g^1</td>
<td>0.1</td>
</tr>
<tr>
<td>g^2</td>
<td>0.4</td>
</tr>
<tr>
<td>g^3</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Daphne Koller
Causal Reasoning

\[ P(I^1) \approx 0.5 \]
\[ P(I^1 \mid i^0) \approx 0.39 \]
\[ P(I^1 \mid i^0, d^0) \approx 0.51 \]
Evidential Reasoning

\[ P(d^1) = 0.4 \]
\[ P(d^1 | g^3) \approx 0.63 \]
\[ P(i^1) = 0.3 \]
\[ P(i^1 | g^3) \approx 0.08 \]

<table>
<thead>
<tr>
<th></th>
<th>( g^1 )</th>
<th>( g^2 )</th>
<th>( g^3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( i^0, d^0 )</td>
<td>0.3</td>
<td>0.4</td>
<td>0.3</td>
</tr>
<tr>
<td>( i^0, d^1 )</td>
<td>0.05</td>
<td>0.25</td>
<td>0.7</td>
</tr>
<tr>
<td>( i^1, d^0 )</td>
<td>0.9</td>
<td>0.08</td>
<td>0.02</td>
</tr>
<tr>
<td>( i^1, d^1 )</td>
<td>0.5</td>
<td>0.3</td>
<td>0.2</td>
</tr>
</tbody>
</table>
Class is hard!

Student gets a C 😞

Intercausal Reasoning

\[ P(d^1) = 0.4 \]
\[ P(d^1 | g^3) \approx 0.63 \]

\[ P(i^1) = 0.3 \]
\[ P(i^1 | g^3) \approx 0.08 \]
\[ P(i^1 | g^3, d^1) \approx 0.11 \]
Intercausal Reasoning Explained

Explaining away

\[ \begin{array}{ccc|c|c}
X_1 & X_2 & Y & \text{Prob} \\
\hline
0 & 0 & 0 & 0.25 \\
0 & 1 & 1 & 0.25 \\
1 & 0 & 1 & 0.25 \\
1 & 1 & 1 & 0.25 \\
\end{array} \]

\[ P(X_{1|2}) = \frac{2}{3}, \quad P(X_{2|1}) = \frac{2}{3} \]

\[ \text{condition } X_1 \quad P(X_{1|2}) = 0.3 \]

Daphne Koller
Intercausal Reasoning II

Class is hard!

Student gets a B 😊

\[ P(i^1) = 0.3 \]
\[ P(i^1 | g^2) \approx 0.175 \]
\[ P(i^1 | g^2, d^1) \approx 0.34 \]
Student Aces the SAT

- What happens to the posterior probability that the class is hard?

Student gets a C 😞

Student aces the SAT 😊
Student Aces the SAT

\[
P(d^1) = 0.4 \\
P(d^1 \mid g^3) \approx 0.63 \\
P(d^1 \mid g^3, s^1) \approx 0.76 \\
P(i^1) = 0.3 \\
P(i^1 \mid g^3) \approx 0.08 \\
P(i^1 \mid g^3, s^1) \approx 0.58
\]

Student aces the SAT 😊

Student gets a C 😞
Probabilistic Graphical Models

Bayesian Networks

Flow of Probabilistic Influence

Representation
When can X influence Y?

- $X \rightarrow Y$ ✓
- $X \leftarrow Y$ ✓
- $X \rightarrow W \rightarrow Y$ ✓
- $X \leftarrow W \leftarrow Y$ ✓
- $X \leftarrow W \rightarrow Y$ ✓
- $X \leftarrow W \leftarrow Y$ ✓
- $X \rightarrow W \leftarrow Y$ ×

Condition on X changes belief about Y

Diagram:

- Difficulty
- Intelligence
- Grade
- Letter
- SAT

Daphne Koller
Active Trails

• A trail $X_1 \ldots X_n$ is active if:

  it has no v-structures $X_{i-1} \rightarrow X_i \leftarrow X_{i+1}$
When can X influence Y
Given evidence about Z

- $X \rightarrow Y$
- $X \leftarrow Y$
- $X \rightarrow W \rightarrow Y$
- $X \leftarrow W \leftarrow Y$
- $X \leftarrow W \rightarrow Y$
- $X \rightarrow W \leftarrow Y$

$W \not\in Z$ | $W \in Z$
---|---

Daphne Koller
When can X influence Y given evidence about Z

- S \rightarrow I \rightarrow G \rightarrow D allows influence to flow when:
  - I is observed
  - I is not observed
  - S is observed
  - S is not observed

Daphne Koller
Active Trails

- A trail $X_1 - \ldots - X_n$ is active given $Z$ if:
  - for any v-structure $X_{i-1} \rightarrow X_i \leftarrow X_{i+1}$ we have that $X_i$ or one of its descendants $\in Z$
  - no other $X_i$ is in $Z$
Probabilistic Graphical Models

Representation
Independencies

Preliminaries
Independence

• For events $\alpha, \beta$, $P \models \alpha \perp \beta$ if:
  - $P(\alpha, \beta) = P(\alpha) \cdot P(\beta)$
  - $P(\alpha | \beta) = P(\alpha)$
  - $P(\beta | \alpha) = P(\beta)$

• For random variables $X, Y$, $P \models X \perp Y$ if:
  - $P(X, Y) = P(X) \cdot P(Y)$
  - $P(X | Y) = P(X)$
  - $P(Y | X) = P(Y)$
Independence

\[
P(I, D) = P(I) \times P(D)
\]

<table>
<thead>
<tr>
<th>I</th>
<th>D</th>
<th>G</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i^0)</td>
<td>(d^0)</td>
<td>(g^1)</td>
<td>0.126</td>
</tr>
<tr>
<td>(i^0)</td>
<td>(d^0)</td>
<td>(g^2)</td>
<td>0.168</td>
</tr>
<tr>
<td>(i^0)</td>
<td>(d^0)</td>
<td>(g^3)</td>
<td>0.126</td>
</tr>
<tr>
<td>(i^0)</td>
<td>(d^1)</td>
<td>(g^1)</td>
<td>0.009</td>
</tr>
<tr>
<td>(i^0)</td>
<td>(d^1)</td>
<td>(g^2)</td>
<td>0.045</td>
</tr>
<tr>
<td>(i^0)</td>
<td>(d^1)</td>
<td>(g^3)</td>
<td>0.126</td>
</tr>
<tr>
<td>(i^1)</td>
<td>(d^0)</td>
<td>(g^1)</td>
<td>0.252</td>
</tr>
<tr>
<td>(i^1)</td>
<td>(d^0)</td>
<td>(g^2)</td>
<td>0.0224</td>
</tr>
<tr>
<td>(i^1)</td>
<td>(d^0)</td>
<td>(g^3)</td>
<td>0.0056</td>
</tr>
<tr>
<td>(i^1)</td>
<td>(d^1)</td>
<td>(g^1)</td>
<td>0.06</td>
</tr>
<tr>
<td>(i^1)</td>
<td>(d^1)</td>
<td>(g^2)</td>
<td>0.036</td>
</tr>
<tr>
<td>(i^1)</td>
<td>(d^1)</td>
<td>(g^3)</td>
<td>0.024</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>I</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i^0)</td>
<td>0.6</td>
</tr>
<tr>
<td>(i^1)</td>
<td>0.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>(d^0)</td>
<td>0.7</td>
</tr>
<tr>
<td>(d^1)</td>
<td>0.3</td>
</tr>
</tbody>
</table>
**Conditional Independence**

- For (sets of) random variables $X, Y, Z$
  
  $P \models (X \perp Y \mid Z)$ if:
  
  - $P(X, Y \mid Z) = P(X \mid Z) \cdot P(Y \mid Z)$
  - $P(X \mid Y, Z) = P(X \mid Z)$
  - $P(Y \mid X, Z) = P(Y \mid Z)$
  - $P(X, Y, Z) \propto \phi_1(X, Z) \cdot \phi_2(Y, Z)$
Conditional Independence

\[ \forall x_1 \perp x_2 \quad \rho \models (x_1 \perp x_2 | c) \]
Conditional Independence

$P(I, S, G)$

<table>
<thead>
<tr>
<th>I</th>
<th>S</th>
<th>G</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i^0$</td>
<td>$s^0$</td>
<td>$g^1$</td>
<td>0.114</td>
</tr>
<tr>
<td>$i^0$</td>
<td>$s^0$</td>
<td>$g^2$</td>
<td>0.1938</td>
</tr>
<tr>
<td>$i^0$</td>
<td>$s^0$</td>
<td>$g^3$</td>
<td>0.2622</td>
</tr>
<tr>
<td>$i^0$</td>
<td>$s^1$</td>
<td>$g^1$</td>
<td>0.006</td>
</tr>
<tr>
<td>$i^0$</td>
<td>$s^1$</td>
<td>$g^2$</td>
<td>0.0102</td>
</tr>
<tr>
<td>$i^0$</td>
<td>$s^1$</td>
<td>$g^3$</td>
<td>0.0138</td>
</tr>
<tr>
<td>$i^1$</td>
<td>$s^0$</td>
<td>$g^1$</td>
<td>0.252</td>
</tr>
<tr>
<td>$i^1$</td>
<td>$s^0$</td>
<td>$g^2$</td>
<td>0.0224</td>
</tr>
<tr>
<td>$i^1$</td>
<td>$s^0$</td>
<td>$g^3$</td>
<td>0.0056</td>
</tr>
<tr>
<td>$i^1$</td>
<td>$s^1$</td>
<td>$g^1$</td>
<td>0.108</td>
</tr>
<tr>
<td>$i^1$</td>
<td>$s^1$</td>
<td>$g^2$</td>
<td>0.0096</td>
</tr>
<tr>
<td>$i^1$</td>
<td>$s^1$</td>
<td>$g^3$</td>
<td>0.0024</td>
</tr>
</tbody>
</table>

$P(S, G | i^0)$

<table>
<thead>
<tr>
<th>S</th>
<th>G</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s^0$</td>
<td>$g^1$</td>
<td>0.19</td>
</tr>
<tr>
<td>$s^0$</td>
<td>$g^2$</td>
<td>0.323</td>
</tr>
<tr>
<td>$s^0$</td>
<td>$g^3$</td>
<td>0.437</td>
</tr>
<tr>
<td>$s^1$</td>
<td>$g^1$</td>
<td>0.01</td>
</tr>
<tr>
<td>$s^1$</td>
<td>$g^2$</td>
<td>0.017</td>
</tr>
<tr>
<td>$s^1$</td>
<td>$g^3$</td>
<td>0.023</td>
</tr>
</tbody>
</table>

$P(S | i^0)$

<table>
<thead>
<tr>
<th>S</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s^0$</td>
<td>0.95</td>
</tr>
<tr>
<td>$s^1$</td>
<td>0.05</td>
</tr>
</tbody>
</table>

$P(G | i^0)$

<table>
<thead>
<tr>
<th>G</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$g^1$</td>
<td>0.2</td>
</tr>
<tr>
<td>$g^2$</td>
<td>0.34</td>
</tr>
<tr>
<td>$g^3$</td>
<td>0.46</td>
</tr>
</tbody>
</table>
Conditioning can Lose Independences

<table>
<thead>
<tr>
<th>I</th>
<th>D</th>
<th>G</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>i₀</td>
<td>d₀</td>
<td>g¹</td>
<td>0.126</td>
</tr>
<tr>
<td>i₀</td>
<td>d₀</td>
<td>g²</td>
<td>0.168</td>
</tr>
<tr>
<td>i₀</td>
<td>d₀</td>
<td>g³</td>
<td>0.126</td>
</tr>
<tr>
<td>i₀</td>
<td>d¹</td>
<td>g¹</td>
<td>0.009</td>
</tr>
<tr>
<td>i₀</td>
<td>d¹</td>
<td>g²</td>
<td>0.045</td>
</tr>
<tr>
<td>i₀</td>
<td>d¹</td>
<td>g³</td>
<td>0.126</td>
</tr>
<tr>
<td>i¹</td>
<td>d₀</td>
<td>g¹</td>
<td>0.252</td>
</tr>
<tr>
<td>i¹</td>
<td>d₀</td>
<td>g²</td>
<td>0.0224</td>
</tr>
<tr>
<td>i¹</td>
<td>d₀</td>
<td>g³</td>
<td>0.0056</td>
</tr>
<tr>
<td>i¹</td>
<td>d¹</td>
<td>g¹</td>
<td>0.06</td>
</tr>
<tr>
<td>i¹</td>
<td>d¹</td>
<td>g²</td>
<td>0.036</td>
</tr>
<tr>
<td>i¹</td>
<td>d¹</td>
<td>g³</td>
<td>0.024</td>
</tr>
</tbody>
</table>

\[ P(I,D \mid g^1) \]

<table>
<thead>
<tr>
<th>I</th>
<th>D</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>i₀</td>
<td>d₀</td>
<td>0.282</td>
</tr>
<tr>
<td>i₀</td>
<td>d¹</td>
<td>0.02</td>
</tr>
<tr>
<td>i¹</td>
<td>d⁰</td>
<td>0.564</td>
</tr>
<tr>
<td>i¹</td>
<td>d¹</td>
<td>0.134</td>
</tr>
</tbody>
</table>
Probabilistic Graphical Models

Bayesian Networks

Representation
Independencies
Independence & Factorization

\[ P(X,Y) = P(X) P(Y) \quad \text{X,Y independent} \]
\[ P(X,Y,Z) \propto \phi_1(X,Z) \phi_2(Y,Z) \quad (X \perp Y \mid Z) \]

• Factorization of a distribution \( P \) implies independencies that hold in \( P \)
• If \( P \) factorizes over \( G \), can we read these independencies from the structure of \( G \)?
Flow of influence & d-separation

**Definition:** $X$ and $Y$ are *d-separated* in $G$ given $Z$ if there is no active trail in $G$ between $X$ and $Y$ given $Z$.

Notation: $d$-sep$_G(X, Y \mid Z)$
Factorization $\Leftrightarrow$ Independence: BNs

**Theorem:** If $P$ factorizes over $G$, and $d$-sep$_G(X, Y \mid Z)$ then $P$ satisfies $(X \perp Y \mid Z)$.

\[
P(D, I, G, S, L) = P(D)P(I)P(G \mid D, I)P(S \mid I)P(L \mid G)
\]

\[
P(D, S) = \sum_{G, I, L} P(D)P(I)P(G \mid D, I)P(S \mid I)P(L \mid G)
\]

\[
= \sum_{I} P(D)P(I)P(S \mid I) \sum_{G} (P(G \mid D, I) \sum_{L} P(L \mid G))
\]

\[
= P(D)\left(\sum_{I} P(I)P(S \mid I)\right)
\]

$P \Rightarrow D \perp S$
Any node is d-separated from its non-descendants given its parents.

If $P$ factorizes over $G$, then in $P$, any variable is independent of its non-descendants given its parents.
I-maps

• d-separation in $G \Rightarrow P$ satisfies corresponding independence statement

$$I(G) = \{ (X \perp Y \mid Z) : d-sep_G(X, Y \mid Z) \}$$

• Definition: If $P$ satisfies $I(G)$, we say that $G$ is an I-map (independency map) of $P$
I-maps

\[
\begin{array}{|c|c|c|}
\hline
I & D & \text{Prob} \\
\hline
i^0 & d^0 & 0.42 \\
i^0 & d^1 & 0.18 \\
i^1 & d^0 & 0.28 \\
i^1 & d^1 & 0.12 \\
\hline
\end{array}
\]

\[
\begin{array}{|c|c|c|}
\hline
I & D & \text{Prob.} \\
\hline
i^0 & d^0 & 0.282 \\
i^0 & d^1 & 0.02 \\
i^1 & d^0 & 0.564 \\
i^1 & d^1 & 0.134 \\
\hline
\end{array}
\]
Factorization $\iff$ Independence: BNs

**Theorem:** If $P$ factorizes over $G$, then $G$ is an I-map for $P$

Can read from $G$ independencies in $P$ regardless of parameters
Independence $\iff$ Factorization

**Theorem:** If $G$ is an I-map for $P$, then $P$ factorizes over $G$.

\[ P(D, I, G, S, L) = P(D)P(I \mid D)P(G \mid D, I)P(S \mid D, I, G)P(L \mid D, I, G, S) \]

\[ P(D, I, G, S, L) = P(D)P(I)P(G \mid D, I)P(S \mid I)P(L \mid G) \]
Summary

Two equivalent views of graph structure:
- Factorization: $G$ allows $P$ to be represented
- I-map: Independencies encoded by $G$ hold in $P$

If $P$ factorizes over a graph $G$, we can read from the graph independencies that must hold in $P$ (an independency map)
Probabilistic Graphical Models

Representation
Bayesian Networks

Naïve Bayes
Naïve Bayes Model

\[ P(C, X_1, \ldots, X_n) = P(C) \prod_{i=1}^{n} P(X_i \mid C) \]

\((X_i \perp X_j \mid C)\) for all \(X_i, X_j\)

\(x_i, x_j\) conditionally independent given \(C\)
Naïve Bayes Classifier

\[
\frac{P(C = c^1 \mid x_1, \ldots, x_n)}{P(C = c^2 \mid x_1, \ldots, x_n)} = \frac{P(C = c^1)}{P(C = c^2)} \prod_{i=1}^{n} \frac{P(x_i \mid C = c^1)}{P(x_i \mid C = c^2)}
\]

Daphne Koller
Bernoulli Naïve Bayes for Text

\[
\frac{P(C = c^1 \mid x_1, \ldots, x_n)}{P(C = c^2 \mid x_1, \ldots, x_n)} = \frac{P(C = c^1)}{P(C = c^2)} \prod_{i=1}^{n} \frac{P(x_i \mid C = c^1)}{P(x_i \mid C = c^2)}
\]

Financial
Pets

<table>
<thead>
<tr>
<th></th>
<th>cat</th>
<th>dog</th>
<th>buy</th>
<th>sell</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.001</td>
<td>0.3</td>
<td>0.4</td>
<td>0.02</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Daphne Koller
Multinomial Naïve Bayes for Text

\[ \frac{P(C = c^1 \mid x_1, \ldots, x_n)}{P(C = c^2 \mid x_1, \ldots, x_n)} = \frac{P(C = c^1)}{P(C = c^2)} \prod_{i=1}^{n} \frac{P(x_i \mid C = c^1)}{P(x_i \mid C = c^2)} \]

Daphne Koller
Summary

• Simple approach for classification
  – Computationally efficient
  – Easy to construct

• Surprisingly effective in domains with many weakly relevant features

• Strong independence assumptions reduce performance when many features are strongly correlated
Probabilistic Graphical Models

Representation
Bayesian Networks

Application: Diagnosis
Medical Diagnosis: Pathfinder (1992)

- Help pathologist diagnose lymph node pathologies (60 different diseases)
- Pathfinder I: Rule-based system
- Pathfinder II used naïve Bayes and got superior performance

Heckerman et al.
Medical Diagnosis: Pathfinder (1992)

- Pathfinder III: Naïve Bayes with better knowledge engineering
- No incorrect zero probabilities
- Better calibration of conditional probabilities
  - $P(\text{finding}_1 \mid \text{disease}_1)$ to $P(\text{finding}_1 \mid \text{disease}_2)$
  - Not $P(\text{finding}_1 \mid \text{disease})$ to $P(\text{finding}_2 \mid \text{disease})$

Heckerman et al.
Medical Diagnosis: Pathfinder (1992)

- Pathfinder IV: Full Bayesian network
  - Removed incorrect independencies
  - Additional parents led to more accurate estimation of probabilities
- BN model agreed with expert panel in 50/53 cases, vs 47/53 for naïve Bayes model
- Accuracy as high as expert that designed the model

Heckerman et al.
M. Pradhan, G. Provan, B. Middleton, M. Henrion, UAI 94

# of parameters: $2^{1000}$ to 133,931,430 to 8254
Medical Diagnosis (Microsoft)

Thanks to: Eric Horvitz, Microsoft Research
Medical Diagnosis (Microsoft)

There are two ways to search for specific information in OnParenting. In Find by Word, type the word(s) you want to find and get a list of titles relevant to that word. Find by Symptom will help you get information about children's symptoms. Here has tips to target your search.

Describe the child
In the drop-down boxes at the right. Relevant information will appear below.

- Localized pain: Can the child localize, or point to, the site of the pain?
  - No, unable to localize
  - Below the navel to the child's left
  - Above the child's navel
  - Either of the child's sides
  - Below the navel to the child's right
  - Above the navel to the child's right
  - Above the navel to the child's left
  - Don't Know

- Age: Toddler
- Sex: Female
- Complaint: Abdominal pain

Results so far

<table>
<thead>
<tr>
<th>Disorder</th>
<th>Relevance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viral gastroenteritis</td>
<td></td>
</tr>
<tr>
<td>Psychosomatic pain</td>
<td></td>
</tr>
<tr>
<td>Urinary tract infection</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
</tr>
</tbody>
</table>

Start Over  Review  Finish

Thanks to: Eric Horvitz, Microsoft Research  Daphne Koller
Fault Diagnosis

- Microsoft troubleshooters
Fault Diagnosis

• Many examples:
  – Microsoft troubleshooters
  – Car repair

• Benefits:
  – Flexible user interface
  – Easy to design and maintain

Daphne Koller