Two applications of Bayesian networks

Jiří Vomlel

Laboratory for Intelligent Systems
University of Economics, Prague

Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic

This presentation is available at:
http://www.utia.cas.cz/vomlel/
Contents:

- Bayesian networks as a model for reasoning with uncertainty
- Building probabilistic models
- Building “good” strategies using the models
- Application 1: Adaptive testing
- Application 2: Decision-theoretic troubleshooting
An example of a Bayesian network:
Building Bayesian network models

three basic approaches

• Discussions with domain experts: expert knowledge is used to get the structure and parameters of the model.

• A dataset of records is collected and a machine learning method is used to construct a model and estimate its parameters.

• A combination of previous two: e.g. experts helps with the structure, data are used to estimate parameters.
An example of a strategy:

$X_3$ is more difficult question than $X_2$ which is more difficult than $X_1$. 
Building strategies using the models

For all terminal nodes $\ell \in \mathcal{L}(s)$ of a strategy $s$ we define:

- steps that were performed to get to that node (e.g. questions answered in a certain way). It is called collected evidence $e_{\ell}$.
- Using the probabilistic model of the domain we can compute probability of getting to a terminal node $P(e_{\ell})$.
- Also during the process, when we have collected certain evidence $e$ we can update the probability of getting to a terminal node, which now corresponds to conditional probability $P(e_{\ell})$. 
Building strategies using the models

For all terminal nodes $\ell \in \mathcal{L}(s)$ of a strategy $s$ we have also defined:

- an evaluation function $f : \bigcup_{s \in S} \mathcal{L}(s) \mapsto \mathbb{R}$.

For each strategy we can compute:

- expected value of the strategy:

\[
E_f(s) = \sum_{\ell \in \mathcal{L}(s)} P(e_\ell) \cdot f(e_\ell)
\]

The goal:

- find a strategy that maximizes (minimizes) its expected value
Using entropy as an information measure

“The lower the entropy of a probability distribution the more we know.”

\[ H (P(S)) = - \sum_{s} P(S = s) \cdot \log P(S = s) \]
Entropy in node $n$

$$H(e_n) = H(P(S \mid e_n))$$

Expected entropy at the end of test $t$

$$E_H(t) = \sum_{\ell \in \mathcal{L}(t)} P(e_\ell) \cdot H(e_\ell)$$

$\mathcal{T}$ ... the set of all possible tests (e.g. of a given length)

A test $t^*$ is optimal iff

$$t^* = \arg \min_{t \in \mathcal{T}} E_H(t).$$
Application 1: Adaptive test of basic operations with fractions

Examples of tasks:

\[ T_1: \quad \left( \frac{3}{4} \cdot \frac{5}{6} \right) - \frac{1}{8} = \frac{15}{24} - \frac{1}{8} = \frac{5}{8} - \frac{1}{8} = \frac{4}{8} = \frac{1}{2} \]

\[ T_2: \quad \frac{1}{6} + \frac{1}{12} = \frac{2}{12} + \frac{1}{12} = \frac{3}{12} = \frac{1}{4} \]

\[ T_3: \quad \frac{1}{4} \cdot 1\frac{1}{2} = \frac{1}{4} \cdot \frac{3}{2} = \frac{3}{8} \]

\[ T_4: \quad \left( \frac{1}{2} \cdot \frac{1}{2} \right) \cdot \left( \frac{1}{3} + \frac{1}{3} \right) = \frac{1}{4} \cdot \frac{2}{3} = \frac{2}{12} = \frac{1}{6} \]
**Elementary and operational skills**

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP</td>
<td>Comparison (common numerator or denominator)</td>
<td>$\frac{1}{2} &gt; \frac{1}{3}$, $\frac{2}{3} &gt; \frac{1}{3}$</td>
</tr>
<tr>
<td>AD</td>
<td>Addition (comm. denom.)</td>
<td>$\frac{1}{7} + \frac{2}{7} = \frac{1+2}{7} = \frac{3}{7}$</td>
</tr>
<tr>
<td>SB</td>
<td>Subtract. (comm. denom.)</td>
<td>$\frac{2}{5} - \frac{1}{5} = \frac{2-1}{5} = \frac{1}{5}$</td>
</tr>
<tr>
<td>MT</td>
<td>Multiplication</td>
<td>$\frac{1}{2} \cdot \frac{3}{5} = \frac{3}{10}$</td>
</tr>
<tr>
<td>CD</td>
<td>Common denominator</td>
<td>$(\frac{1}{2}, \frac{2}{3}) = (\frac{3}{6}, \frac{4}{6})$</td>
</tr>
<tr>
<td>CL</td>
<td>Cancelling out</td>
<td>$\frac{4}{6} = \frac{2\cdot2}{2\cdot3} = \frac{2}{3}$</td>
</tr>
<tr>
<td>CIM</td>
<td>Conv. to mixed numbers</td>
<td>$\frac{7}{2} = \frac{3\cdot2+1}{2} = 3\frac{1}{2}$</td>
</tr>
<tr>
<td>CMI</td>
<td>Conv. to improp. fractions</td>
<td>$3\frac{1}{2} = \frac{3\cdot2+1}{2} = \frac{7}{2}$</td>
</tr>
</tbody>
</table>
# Misconceptions

<table>
<thead>
<tr>
<th>Label</th>
<th>Description</th>
<th>Occurrence</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAD</td>
<td>$\frac{a}{b} + \frac{c}{d} = \frac{a+c}{b+d}$</td>
<td>14.8%</td>
</tr>
<tr>
<td>MSB</td>
<td>$\frac{a}{b} - \frac{c}{d} = \frac{a-c}{b-d}$</td>
<td>9.4%</td>
</tr>
<tr>
<td>MMT1</td>
<td>$\frac{a}{b} \cdot \frac{c}{b} = \frac{a \cdot c}{b}$</td>
<td>14.1%</td>
</tr>
<tr>
<td>MMT2</td>
<td>$\frac{a}{b} \cdot \frac{c}{b} = \frac{a+c}{b \cdot b}$</td>
<td>8.1%</td>
</tr>
<tr>
<td>MMT3</td>
<td>$\frac{a}{b} \cdot \frac{c}{d} = \frac{a \cdot d}{b \cdot c}$</td>
<td>15.4%</td>
</tr>
<tr>
<td>MMT4</td>
<td>$\frac{a}{b} \cdot \frac{c}{d} = \frac{a \cdot c}{b+d}$</td>
<td>8.1%</td>
</tr>
<tr>
<td>MC</td>
<td>$a\frac{b}{c} = \frac{a \cdot b}{c}$</td>
<td>4.0%</td>
</tr>
</tbody>
</table>
Student model
Evidence model for task $T1$

$$\left(\frac{3}{4} \cdot \frac{5}{6}\right) - \frac{1}{8} = \frac{15}{24} - \frac{1}{8} = \frac{5}{8} - \frac{1}{8} = \frac{4}{8} = \frac{1}{2}$$

$T1 \Leftrightarrow MT \& CL \& ACL \& SB \& \neg MMT3 \& \neg MMT4 \& \neg MSB$
**Application 2: Troubleshooting - Light print problem**

- **Problems:** $F_1$ Distribution problem, $F_2$ Defective toner, $F_3$ Corrupted dataflow, and $F_4$ Wrong driver setting.
- **Actions:** $A_1$ Remove, shake and reseat toner, $A_2$ Try another toner, and $A_3$ Cycle power.
- **Questions:** $Q_1$ Is the configuration page printed light?
The task is to find a strategy \( s \in S \) minimising expected cost of repair

\[
E_{CR}(s) = \sum_{\ell \in \mathcal{L}(s)} P(e_{\ell}) \cdot (t(e_{\ell}) + c(e_{\ell})) .
\]
Going commercial...

- **Hugin Expert A/S.**
  software product: Hugin - a Bayesian network tool.

- **Educational Testing Service (ETS)**
  the world’s largest private educational testing organization
  In 2000/2001 more than 3 millions students took the ETS’s largest exam SAT. Research unit doing research on adaptive test using Bayesian networks: [http://www.ets.org/research/](http://www.ets.org/research/)

- **SACSO Project**
  *Systems for Automatic Customer Support Operations*
  - research project of Hewlett Packard and Aalborg University.
  The troubleshooter offered as DezisionWorks by Dezide Ltd.
  [http://www.dezide.com/](http://www.dezide.com/)