Expert Systems
A knowledge-based approach to intelligent systems

Peter Lucas
Department of Information and Knowledge Systems
Institute for Computing and Information Sciences
University of Nijmegen, The Netherlands

Motivation

Intelligent System

Machine Learning

- no data
- small datasets
- missing data
- large search space
⇒ knowledge-based approach (i.e., via knowledge acquisition)

Knowledge Acquisition

Terminology

- Expert system
- Knowledge-based system
- Knowledge system
- Intelligent system
- Intelligent agent

Sometimes used as synonyms, sometimes used to stress differences w.r.t.:
- Acquisition of knowledge (data or human expertise)
- Amount of expertise (expert or not)
- Content of system versus behaviour
- Architecture of system

Approaches & Ingredients

Approach: knowledge modelling at different levels (get a grip on the knowledge):

- Problem-solving method (PSM):
  - diagnostic PSM
  - planning and scheduling PSM
  - design and configuration PSM
  - decision-making PSM
⇒ implemented in a reasoning method

- Knowledge base
⇒ specified in a knowledge-representation formalism
Logical Approach

- Knowledge base (KB) **Horn clauses**:
  \[ \forall x_1 \cdots \forall x_m ((A_1 \land \cdots \land A_n) \rightarrow B) \]
- PSMs with findings \( F \) and solution \( S \):
  - **Deductive solution** (\( S \) follows from KB and \( F \)):
    \[ KB \cup F \models S \]
    and \( KB \cup F \not\models \bot \).
  - **Abductive/inductive solution** (\( S \) explains \( F \)):
    \[ KB \cup S \cup K \models F \]
    where \( K \) stands for contextual knowledge.
  - **Consistency-based solution** (\( S \) and \( F \) are consistent):
    \[ KB \cup S \cup F \not\models \bot \]

Example: Model-based Diagnosis

- Model: representation of normal or abnormal behaviour, possibly also of the internal structure
- Formalisation:
  - consistency-based diagnosis, and
  - abductive diagnosis

Consistency-based Diagnosis

Discrepancy between predicted behaviour and observed behaviour \( \Rightarrow \) fault (defect)!


Normal Behaviour

- **SD (System Description)**:
  \[ MUL(M_1), MUL(M_2), MUL(M_3), \]
  \[ ADD(A_1), ADD(A_2) \]
  \[ in_1(A_1) = out(M_1), in_2(A_1) = out(M_2) \]
  \[ in_1(A_2) = out(M_2), in_2(A_2) = out(M_3) \]
  \[ \forall x (MUL(x) \rightarrow in_1(x) \times in_2(x) = out(x)) \]
  \[ \forall x (ADD(x) \rightarrow in_1(x) + in_2(x) = out(x)) \]
- **COMPS** = \{\( M_1, M_2, M_3, A_1, A_2 \)\}
**AB Predicate**

- $AB(c)$: component $c$ is abnormal
- $\neg AB(c)$: component $c$ is normal

Example (Inverter $I$):

\[
\begin{array}{c}
1 \\
\end{array}
\xrightarrow{\text{I}}
\begin{array}{c}
0 \\
\end{array}
\]

\[SD = \{\forall x((\text{INV}(x) \land \neg AB(x)) \rightarrow \neg (\text{out}(x) = \text{in}(x))), \text{INV}(I)\}\]

- Input: in($I$) = 1
- Observed output: out($I$) = 1

SD $\cup \{\text{in}(I) = 1, \text{out}(I) = 1\} \cup \{\neg \text{AB}(I)\} \models \bot$

(assumption that $I$ is normal is inconsistent)

SD $\cup \{\text{in}(I) = 1, \text{out}(I) = 1\} \cup \{\text{AB}(I)\} \not\models \bot$

(assumption that $I$ is abnormal is consistent)

**Abnormality Assumptions**

\[\text{SYS} = (\text{SD}, \text{COMPS}):\]

- SD (System Description):
  \[
  \forall x((\text{MUL}(x) \land \neg \text{AB}(x)) \rightarrow \text{in}_1(x) \times \text{in}_2(x) = \text{out}(x))
  \]
  \[
  \forall x((\text{ADD}(x) \land \neg \text{AB}(x)) \rightarrow \text{in}_1(x) + \text{in}_2(x) = \text{out}(x))
  \]

- (Ab)normality assumptions $D = \{\neg \text{Ab}(c) | c \in \text{COMPS} - \Delta\} \cup \{\text{Ab}(c) | c \in \Delta\}$, $\Delta \subseteq \{M_1, M_2, M_3, A_1, A_2\}$

**Which Components are Faulty?**

Possible diagnoses (faulty componenten) $\Delta$:

- $\Delta = \{A_1\}, \{M_1\}, \{M_2, M_3\}, \{A_2, M_2\}$, since
  \[
  \text{SD} \cup D \cup \text{OBS} \not\models \bot
  \]
  where $D = \{\neg \text{Ab}(c) | c \in \text{COMPS} - \Delta\} \cup \{\text{Ab}(c) | c \in D\}$

- $\Delta$ is always a smallest set, since $\Delta = \text{COMPS}$ would also be a diagnosis otherwise

**Abductive Diagnosis**

Correspondence between predicted abnormal behaviour and observed behaviour $\Rightarrow$ defect!

Originator:

### Causal Models

- **Causality**: combination of Causes have particular Effects
- **Logical representation**: 
  \[ \text{Cause}_1 \land \cdots \land \text{Cause}_n \rightarrow \text{Effect} \]
- **Example**: \( \text{fever} \rightarrow \text{chills} \)

### Weak and Strong Causality

- **Strong causality**: \( C \rightarrow E \)
  “If \( C \) is present, then \( E \) must be present as well”
- **Weak causality**: \( C \land \alpha \rightarrow E \)
  “If \( C \) is present, then \( E \) may be present as well” (\( \alpha \): incompleteness assumption)

### Prediction

- **Causal specification**: \( \Sigma = (\Delta, \Phi, \mathcal{R}) \), with:
  - \( \Delta \): possible causes and incompleteness assumptions
  - \( \Phi \): observable facts
  - \( \mathcal{R} \): causal model
- **Prediction \( V \subseteq \Delta \)**:
  \[ \mathcal{R} \cup V \vdash E \]
  with \( E \subseteq \Phi \) (\( E \) is observable)

### Diagnostic Problem

- **Causal specification**: \( \Sigma = (\Delta, \Phi, \mathcal{R}) \)
- **(Actually) observed facts**: \( F \), for example \( F = \{ \text{myalgia, thirst} \} \)
- **Diagnosis \( D \)**:
  1. Prediction which explains \( F \): \[ \mathcal{R} \cup D \vdash F \]
  2. ... but should not explain too much
Diagnostic Problem

- Causal specification: $\Sigma = (\Delta, \Phi, R)$
- (Actually) observed facts: $F$, for example $F = \{\text{myalgia, thirst}\}$
- Examples of diagnoses:
  $D = \{\text{flu, } \alpha_2\}$, $D' = \{\text{sport, flu}\}$,
  and is $D'' = \{\text{flu, } \alpha_1\}$ a diagnosis?

Consistency Condition

- Causal specification: $\Sigma = (\Delta, \Phi, R)$
- Observed facts: $F = \{\text{myalgia, thirst}\}$
- Facts which should not be explained: $C = \{\neg \text{chills}\}$
- Formally: $D \subseteq \Delta$ is a diagnosis, iff:
  1. $R \cup D \models F$ (covering condition)
  2. $R \cup D \cup C \not\models \bot$ (consistency condition)

Abduction = Anticausal Reasoning

Abduction:

- **Effect, Cause $\rightarrow$ Effect**
- **Cause**

Idea: Reversal of causal relationship

Example:

- $\text{fever} \rightarrow \text{thirst}$

results in:

- $\text{thirst} \rightarrow \text{fever}$

Now:

- $\{\text{thirst} \rightarrow \text{fever}, \text{thirst}\} \models \text{fever}$

Conclusion:

Abduction = deduction with implication reversal
Cost-based Abduction

Express likelihood by means of a cost function:
\[ c : \mathcal{P}(\Delta) \to \mathbb{R} \]

\[ c(D) = \sum_{d \in D} c(d) \]

\( D \subseteq \Delta \) is a *diagnosis* with cost \( c(D) \), iff:

1. \( \mathcal{R} \cup D \models F \) (covering condition)
2. \( \mathcal{R} \cup D \cup C \not\models \perp \) (consistency condition)

Eugene Charniak: cost function \( c \) equal to \(-\log\), then 1–1 mapping cost-based abduction to Bayesian networks

Manual Construction of Bayesian Networks

Qualitative modelling:

- Colonisation by bacterium \( A \)
- Colonisation by bacterium \( B \)
- Colonisation by bacterium \( C \)

Body response to \( A \)
Body response to \( B \)
Body response to \( C \)

Infection

Fever
WBC
ESR

People become colonised by bacteria when entering a hospital, which may give rise to infection

Bayesian-network Modelling

<table>
<thead>
<tr>
<th>Qualitative</th>
<th>Quantitative</th>
</tr>
</thead>
<tbody>
<tr>
<td>causal modelling</td>
<td>interaction modelling</td>
</tr>
</tbody>
</table>

Cause \( \rightarrow \) Effect

<table>
<thead>
<tr>
<th>( \text{Inf} )</th>
<th>( \text{BR}_A )</th>
<th>( \text{BR}_B )</th>
<th>( \text{BR}_C )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t )</td>
<td>( t )</td>
<td>( t )</td>
<td>( t )</td>
</tr>
<tr>
<td>( f )</td>
<td>( f )</td>
<td>( f )</td>
<td>( f )</td>
</tr>
</tbody>
</table>

\[ \text{Pr}(\text{Inf} \mid \text{BR}_A, \text{BR}_B, \text{BR}_C) \]

<table>
<thead>
<tr>
<th>Inf</th>
<th>( \text{BR}_C )</th>
<th>( \text{BR}_C )</th>
<th>( \text{BR}_C )</th>
<th>( \text{BR}_C )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t )</td>
<td>0.8</td>
<td>0.6</td>
<td>0.5</td>
<td>0.3</td>
</tr>
<tr>
<td>( f )</td>
<td>0.2</td>
<td>0.4</td>
<td>0.5</td>
<td>0.7</td>
</tr>
</tbody>
</table>

Problem Solving

As logic, Bayesian networks are declarative, i.e.:

- mathematical basis
- problem to be solved determined by (1) entered findings \( F \) (may include decisions); (2) given hypothesis \( H \):

\[ \text{Pr}(H \mid F) \]

(cf. KB \& F \models H)

Examples:

- Classification and diagnosis: \( D = \arg \max_H \text{Pr}(H \mid F) \)
- Temporal reasoning, prediction, what-if scenarios
- Decision-making based on decision theory

\[ \text{MEU}(D \mid F) = \max_{d \in D} \sum_{x \in X_{\pi(v)}} u(x) \text{Pr}(x \mid d, F) \]
Conclusions

- Knowledge-based approach: need for handles for knowledge modelling
- Model-based approaches support using detailed qualitative models
- Logic can be replaced by set-theoretical or algebraic methods
- Interesting relationships between probabilistic reasoning and qualitative reasoning in model-based systems (e.g., cost-based abduction)