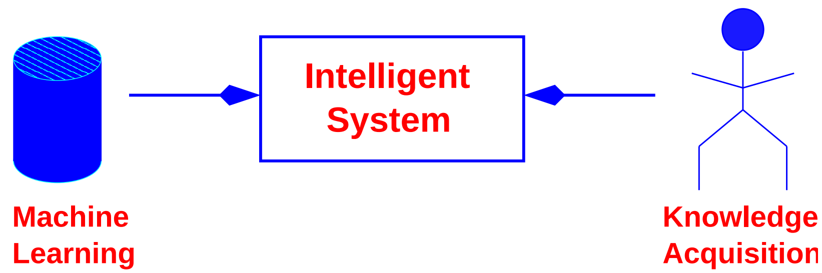
Expert Systems

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



- no data
- small datasets
- missing data
- large search space

⇒ knowledge-based approach
(i.e., via 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**:

  $\forall x_1 \cdots \forall x_m ((A_1 \wedge \cdots \wedge A_n) \rightarrow B)$

- PSMs with findings F and solution S:
  - **Deductive solution** (S follows from KB and F):

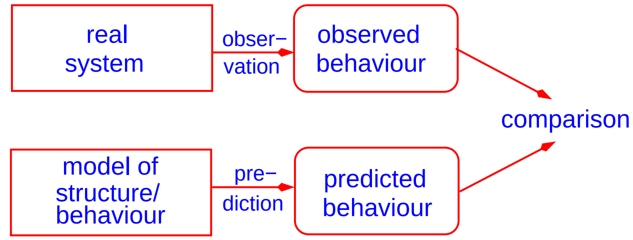
    $\text{KB} \cup F \models S$

    and $\text{KB} \cup F \not\models \bot$.
  - **Abductive/inductive solution** (S *explains* F):

    $\text{KB} \cup S \cup K \models F$

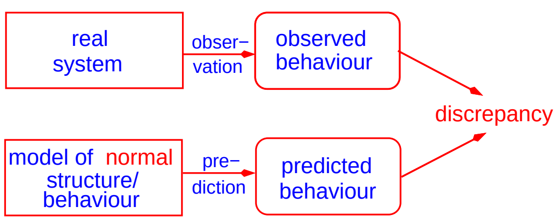
    where K stands for *contextual knowledge*.
  - **Consistency-based solution** (S and F are consistent): $\text{KB} \cup S \cup F \not\models \bot$

Example: Model-based Diagnosis



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

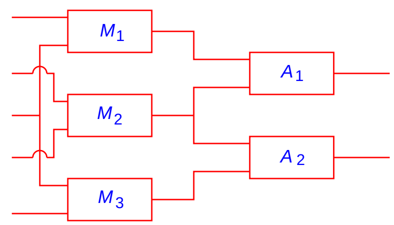
Consistency-based Diagnosis



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

- R. Reiter, "A Theory of diagnosis from first principles", *Artificial Intelligence*, vol. 32, 57–95, 1987.
- J. de Kleer, A.K. Macworth, and R. Reiter, "Characterising diagnoses and systems", *Artificial Intelligence*, vol. 52, 197–222, 1992.

Normal Behaviour



SYStem specification
SYS = (SD, COMPS):

- SD (System Description):

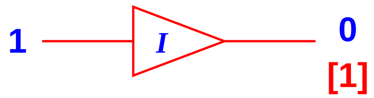
  MUL(M_1), MUL(M_2), MUL(M_3),
  ADD(A_1), ADD(A_2)
  $\text{in}_1(A_1) = \text{out}(M_1), \text{in}_2(A_1) = \text{out}(M_2)$
  $\text{in}_1(A_2) = \text{out}(M_2), \text{in}_2(A_2) = \text{out}(M_3)$
  $\forall x(\text{MUL}(x) \rightarrow \text{in}_1(x) \times \text{in}_2(x) = \text{out}(x))$
  $\forall x(\text{ADD}(x) \rightarrow \text{in}_1(x) + \text{in}_2(x) = \text{out}(x))$

- 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):



$\text{SD} = \{\forall x((\text{INV}(x) \wedge \neg\text{AB}(x)) \rightarrow \neg(\text{out}(x) = \text{in}(x))), \text{INV}(I)\}$

- Input: $\text{in}(I) = 1$
- Observed output: $\text{out}(I) = 1\}$

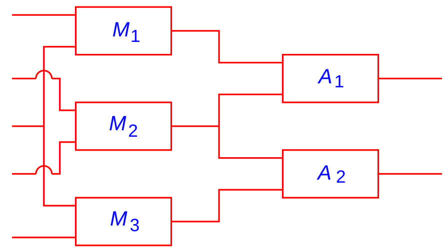
$\text{SD} \cup \{\text{in}(I) = 1, \text{out}(I) = 1\} \cup \{\neg\text{AB}(I)\} \vDash \bot$

(assumption that I is *normal* is *inconsistent*)

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

(assumption that I is *abnormal* is *consistent*)

Abnormality Assumptions



SYStem specification $\text{SYS} = (\text{SD}, \text{COMPS})$:

- SD (System Description):

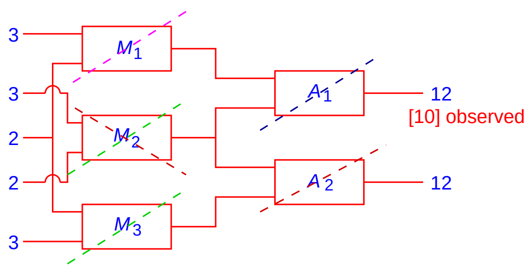
$\forall x((\text{MUL}(x) \wedge \neg\text{Ab}(x)) \rightarrow \text{in}_1(x) \times \text{in}_2(x) = \text{out}(x))$

$\forall x((\text{ADD}(x) \wedge \neg\text{Ab}(x)) \rightarrow \text{in}_1(x) + \text{in}_2(x) = \text{out}(x))$

$\cdots$

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

Which Components are Faulty?



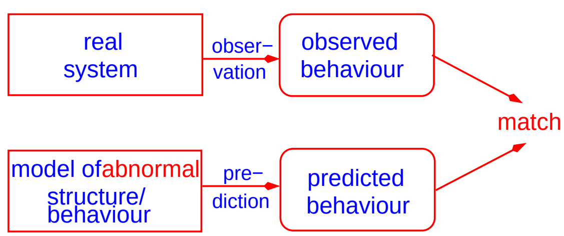
Possible **diagnoses** (faulty componenten) Δ:

- $\Delta = \{A_1\}, \{M_1\}, \{M_2, M_3\}, \{A_2, M_2\}$, since

  $\text{SD} \cup D \cup \text{OBS} \nvDash \bot$

  where $D = \{\neg\text{Ab}(c) \mid c \in \text{COMPS} - \Delta\} \cup \{\text{Ab}(c) \mid c \in D\}$
- Δ is always a *smallest* set, since $\Delta = \text{COMPS}$ would also be a diagnosis otherwise

Abductive Diagnosis

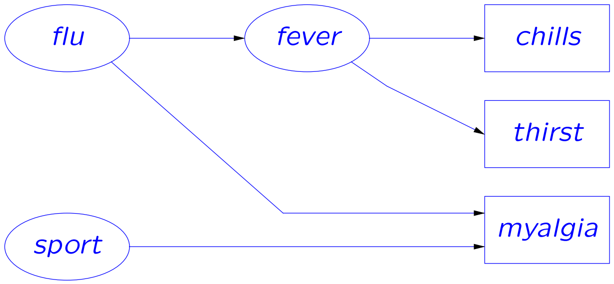


Correspondence between predicted *abnormal* behaviour and observed behaviour $\Rightarrow$ defect!

Originator:

- L. Console, D. Theseider Dupré and P. Torasso, "A theory of diagnosis for incomplete causal models", In: *IJCAI'89*, 1311–1317, 1989

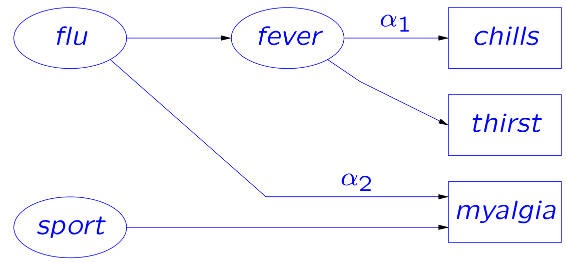
Causal Models



- **Causality**:
  combination of Causes have particular Effects
- **Logical representation**:

  $\text{Cause}_1 \wedge \cdots \wedge \text{Cause}_n \rightarrow \text{Effect}$
- **Example:** *fever* $\rightarrow$ *chills*

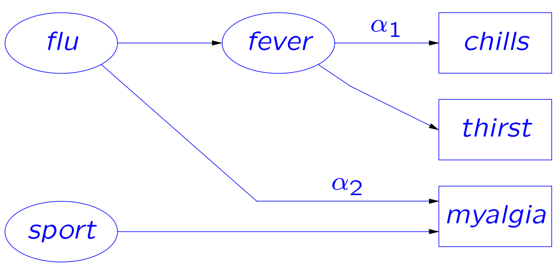
Weak and Strong Causality



Example:
fever $\wedge\ \alpha_1 \rightarrow$ *chills*
fever $\rightarrow$ *thirst*
sport $\rightarrow$ *myalgia*

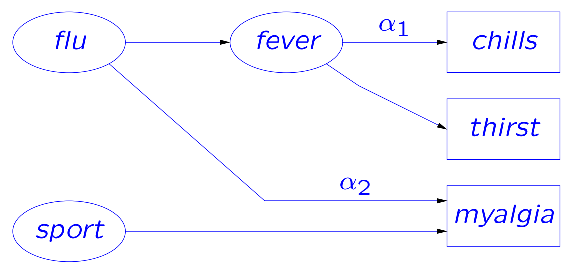
- **Strong causality**: $C \rightarrow E$
  "If C is present, then E *must* be present as well"
- **Weak causality**: $C \wedge \alpha \rightarrow E$
  "If C is present, then E *may* be present as well" (α: incompleteness assumption)

Prediction



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

Diagnostic Problem



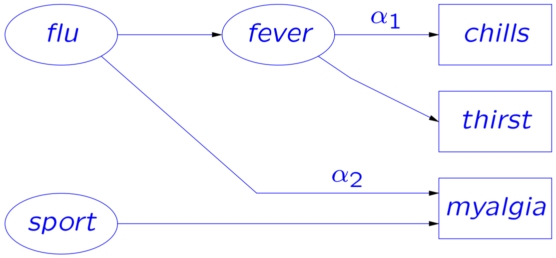
Diagnostic problem $\mathcal{P} = (\Sigma, F)$, with:

- Causal specification: $\Sigma = (\Delta, \Phi, \mathcal{R})$
- (Actually) observed facts: F, for example $F = \{$*myalgia*, *thirst*$\}$
- **Diagnosis** D?

  (1) Prediction which *explains* F: $\boxed{\mathcal{R} \cup D \vDash F}$

  (2) $\cdots$ but should not explain too much

Diagnostic Problem



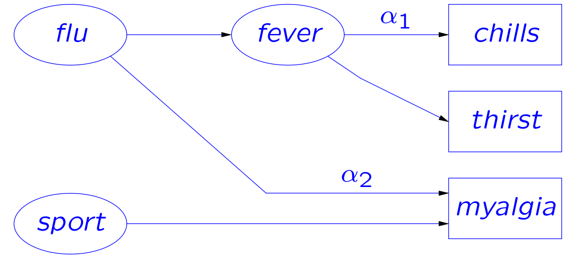
Diagnostic problem $\mathcal{P} = (\Sigma, F)$, with:

- Causal specification: $\Sigma = (\Delta, \Phi, \mathcal{R})$
- (Actually) observed facts: F, for example $F = \{\textit{myalgia}, \textit{thirst}\}$

Examples of diagnoses:

$D = \{\textit{flu}, \alpha_2\}$, $D' = \{\textit{sport}, \textit{flu}\}$,
and is $D'' = \{\textit{flu}, \alpha_1\}$ a diagnosis?

Don't Explain too Much

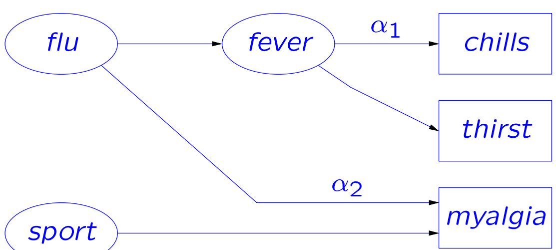


- Causal specification: $\Sigma = (\Delta, \Phi, \mathcal{R})$
- Observed facts: $F = \{\textit{myalgia}, \textit{thirst}\}$
- Facts which should *not* be explained: $C = \{\neg \textit{chills}\}$
- Formally: $D \subseteq \Delta$ is a *diagnosis*, iff:

(1) $\mathcal{R} \cup D \models F$ (covering condition)

(2) $\mathcal{R} \cup D \cup C \not\models \bot$ (consistency condition)

Consistency Condition



- Causal specification: $\Sigma = (\Delta, \Phi, \mathcal{R})$
- Observed facts: $F = \{\textit{myalgia}, \textit{thirst}\}$
- Facts which should *not* be explained:

$C = \{\neg \textit{chills}\}$

$\mathcal{R} \cup \{\textit{flu}, \alpha_1, \alpha_2\} \cup \{\neg \textit{chills}\} \models \bot$

$\Rightarrow D = \{\textit{flu}, \alpha_1, \alpha_2\}$ is *no* diagnosis

Abduction = Anticausal Reasoning

Abduction:

$$\frac{\text{Effect}, \text{Cause} \rightarrow \text{Effect}}{\text{Cause}}$$

Idea: *Reversal of causal relationship*

Example:

fever $\rightarrow$ *thirst*

results in:

thirst $\rightarrow$ *fever*

Now:

$\{\textit{thirst} \rightarrow \textit{fever}, \textit{thirst}\} \models \textit{fever}$

Conclusion:

Abduction = deduction with implication reversal

Cost-based Abduction

Express likelihood by means of a cost function:

$c : \wp(\Delta) \to \mathbb{R}$

often:

$$c(D) = \sum_{d \in D} c(\{d\})$$

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

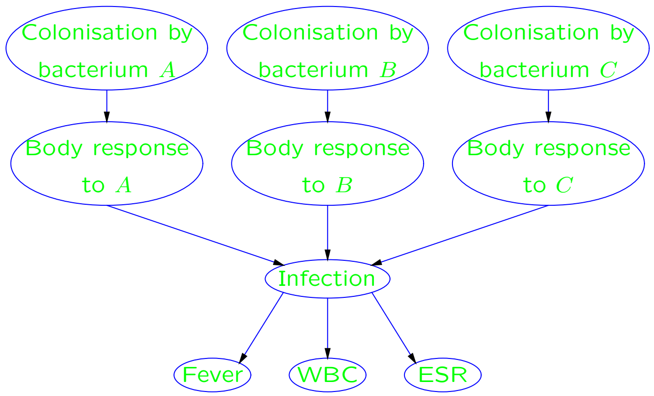
(1) $\mathcal{R} \cup D \vDash F$ (covering condition)

(2) $\mathcal{R} \cup D \cup C \nvDash \bot$ (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:



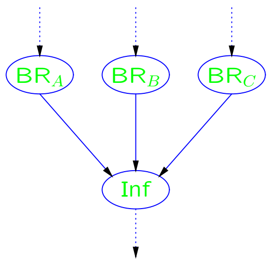
People become colonised by bacteria when entering a hospital, which may give rise to infection

Bayesian-network Modelling

Qualitative

causal modelling

Cause $\to$ Effect



Quantitative

interaction modelling

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

	BR_A							
	t				f			
	BR_B				BR_B			
	t		f		t		f	
	BR_C		BR_C		BR_C		BR_C	
Inf	t	f	t	f	t	f	t	f
t	0.8	0.6	0.5	0.3	0.4	0.2	0.3	0.1
f	0.2	0.4	0.5	0.7	0.6	0.8	0.7	0.9

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:

$\Pr(H \mid F)$ (cf. $\text{KB} \wedge F \vDash H$)

Examples:

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

$$\text{MEU}(D \mid F) = \max_{d \in D} \sum_{x \in X_{\pi(U)}} u(x) \Pr(x \mid d, F)$$

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)