The Model-based Approach to Computer-aided Medical Decision Support

Lecture 3: Building Bayesian Networks

Peter Lucas

`peterl@cs.ru.nl`

Institute for Computing and Information Sciences
Radboud University Nijmegen
The Netherlands

Suitability Bayesian Networks

- Modelling objective, e.g.,
  - uncertain does or does not play a signicant role;
  - white or black box approach: is there sufficient knowledge about the domain?
- Availability of domain experts and data
- Sufficient time available?
- Complexity of the problem: is it decomposable?

Problem Solving

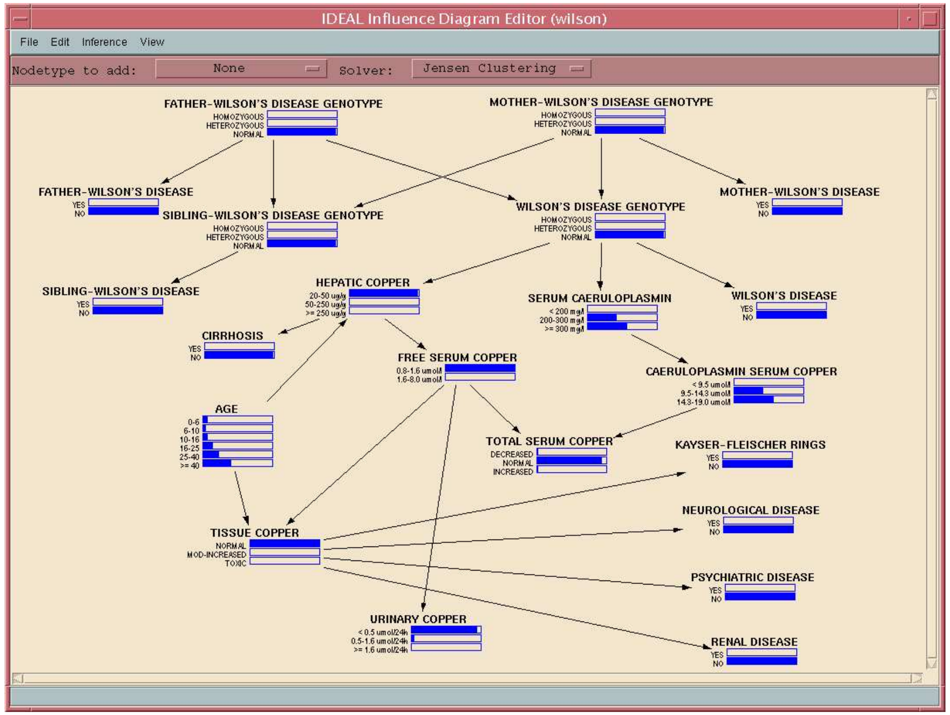
Bayesian networks: a declarative knowledge-representation formalism, i.e.:

- mathematical basis
- problem to be solved determined by (1) entered evidence e (including potential decisions); (2) given hypothesis h: $P(h \mid e)$

Examples:

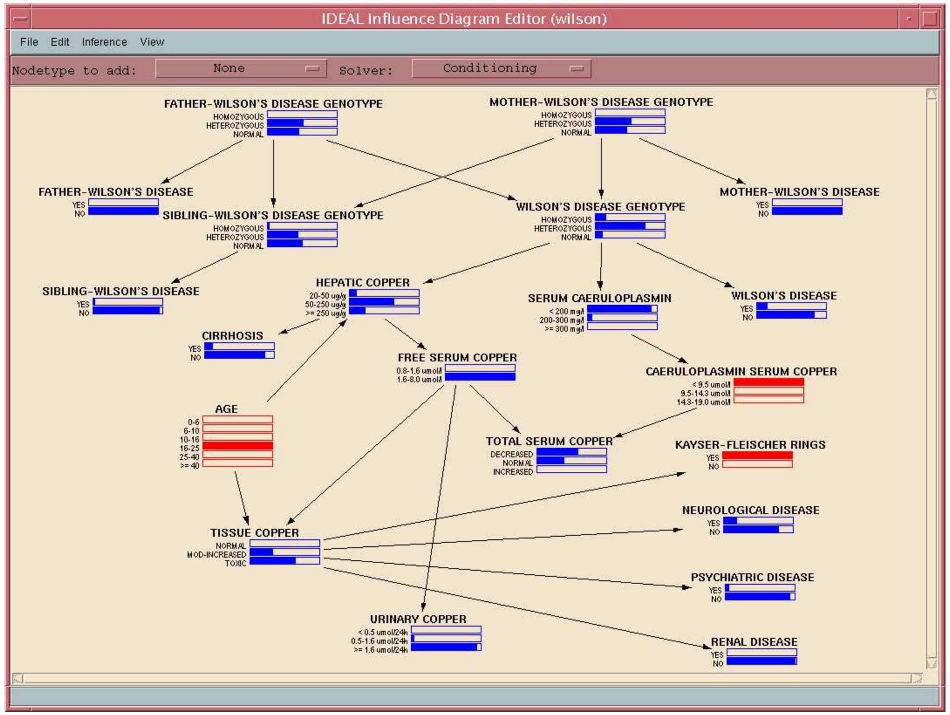
- Description of population (i.e., distributions)
- Classification and diagnosis hypothesis h with maximum $P(h \mid e)$
- Prediction (time dimension)
- Decision making based on *what-if* scenario's

Prior Information



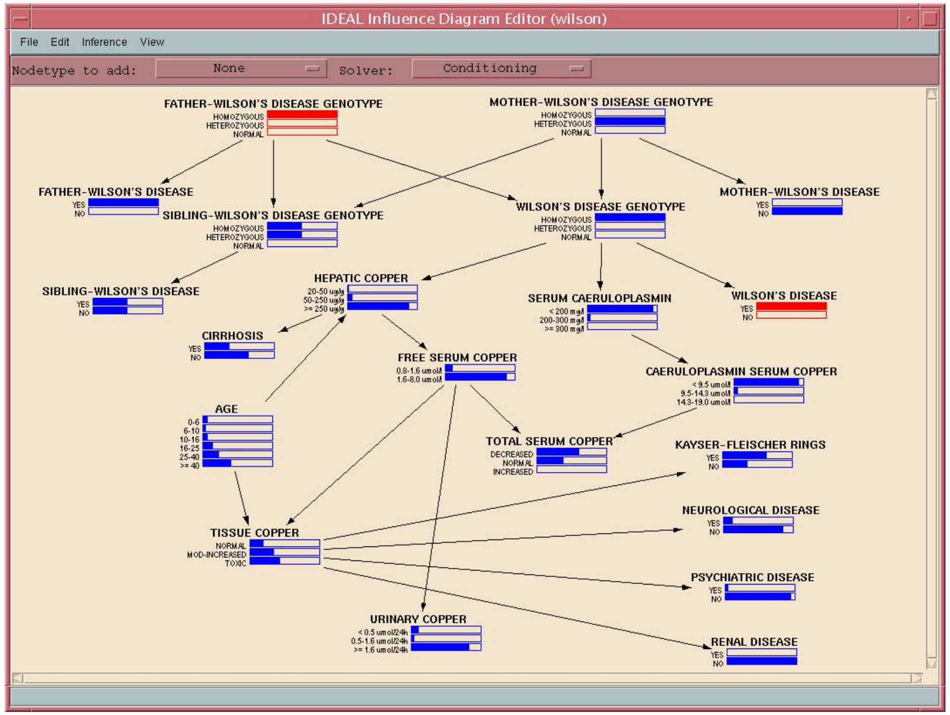
- Marginal probabilities $P(V)$ for every vertex V, e.g., $P(\text{WILSON'S DISEASE} = \textit{yes})$
- Gives description of the population on which the assessed probabilities are based

Diagnostic Problem Solving



- Marginal probabilities $P^*(V) = P(V \mid \mathcal{E})$ for every vertex V, e.g., P(WILSON'S DISEASE $=$ *yes* $\mid \mathcal{E}$) for entered evidence $\mathcal{E}$ (red vertices)
- Gives description of the *sub*population of the original population or individual cases

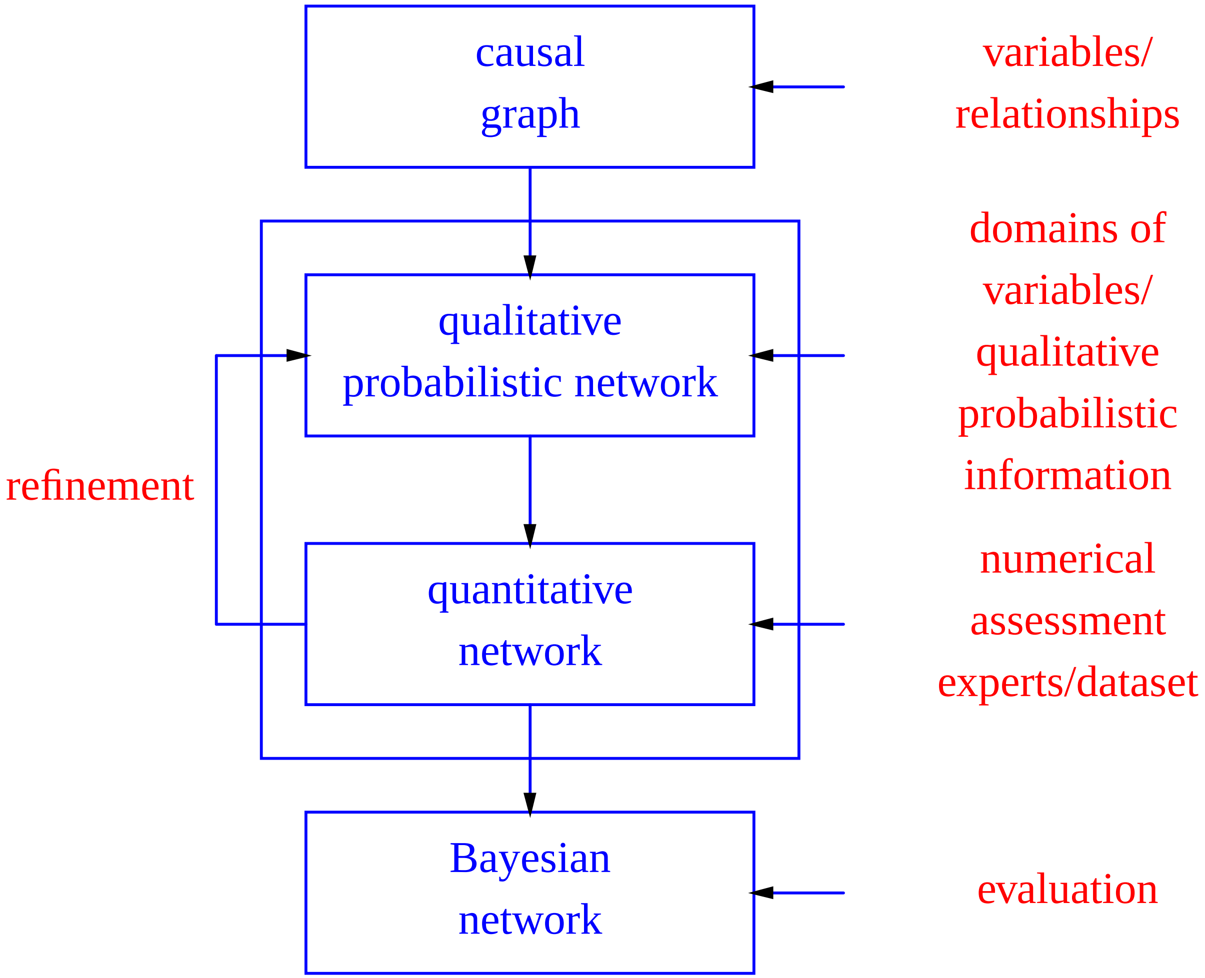
Prediction Associated Findings



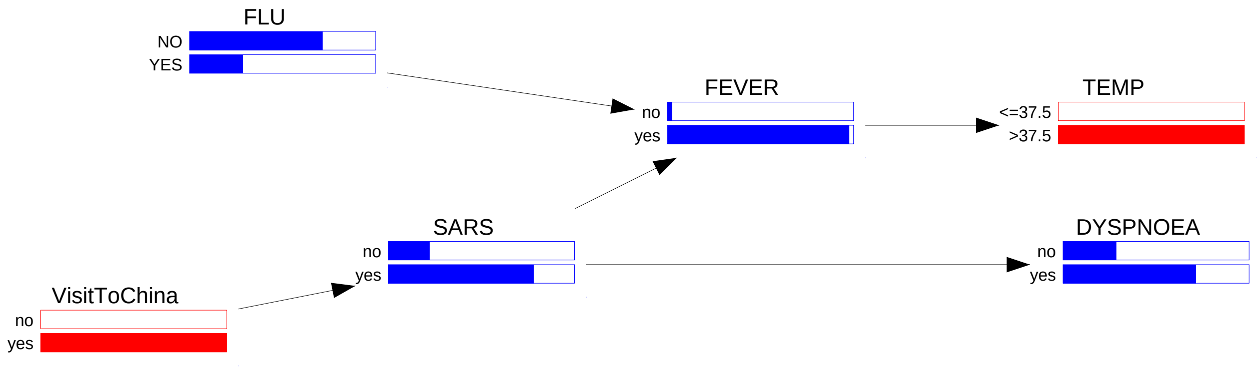
- Marginal probabilities $P^*(V) = P(V \mid \mathcal{E})$, e.g., $P(\text{Kayer-Fleischer Rings} = \textit{yes} \mid \mathcal{E})$ with $\mathcal{E}$ evidence
- Gives description of the findings associated with a given class or category, such as Wilson's disease

Design of Bayesian Network

- Principle: start modelling qualitatively

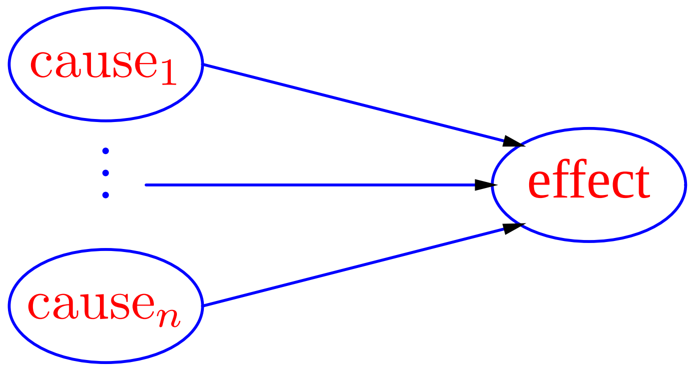


Terminology



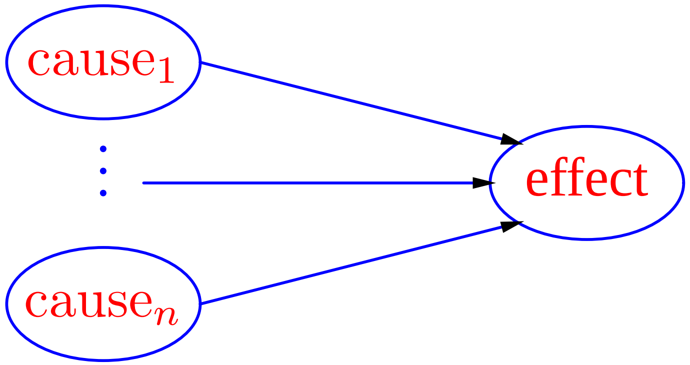
- Parent SARS of child FEVER
- SARS is ancestor of TEMP
- DYSPNOEA is descendant of VISITTOCHINA
- Query node, e.g., FEVER
- Evidence, e.g., VISITTOCHINA and TEMP
- Markov blanket, e.g.,
  for SARS:
  {VISITTOCHINA,DYSPNOEA,FEVER,FLU}

Causal graph: Topology



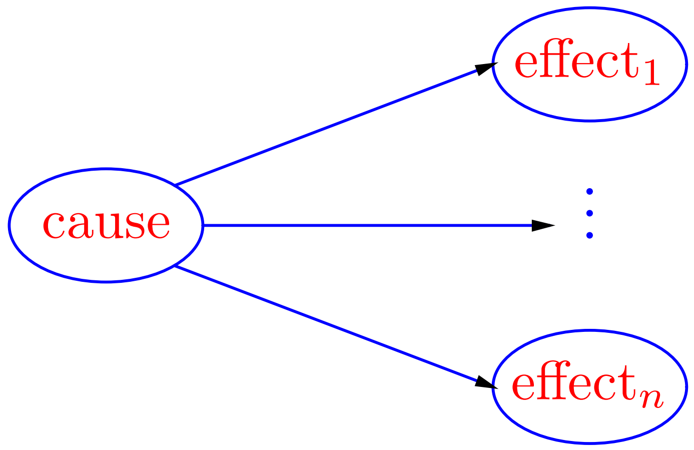
- Identify factors that are relevant
- Determine how those factors are causally related to each other
- The arc $\text{cause} \rightarrow \text{effect}$ does mean that cause is a factor involved in causing effect

Causal graph: Common Effects



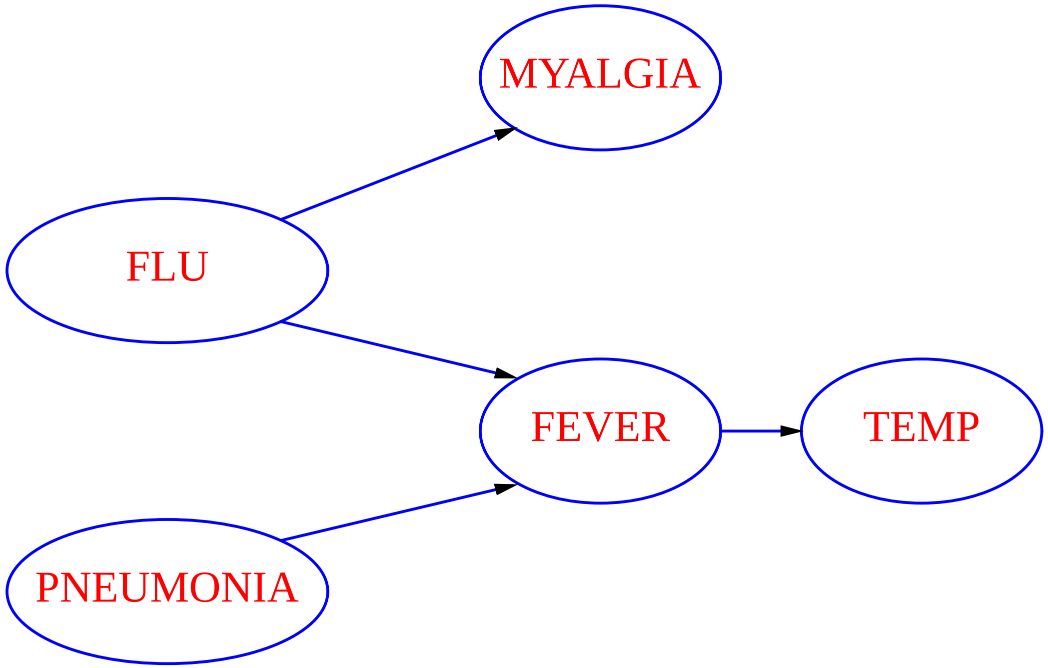
- An effect that has two or more ingoing arcs from other vertices is a common effect of those causes
- Kinds of causal interaction
  - Synergy: POLUTION $\longrightarrow$ CANCER $\longleftarrow$ SMOKING
  - Prevention: VACCINE $\longrightarrow$ DEATH $\longleftarrow$ SMALLPOX
  - XOR: ALKALI $\longrightarrow$ DEATH $\longleftarrow$ ACID

Causal graph: Common Causes



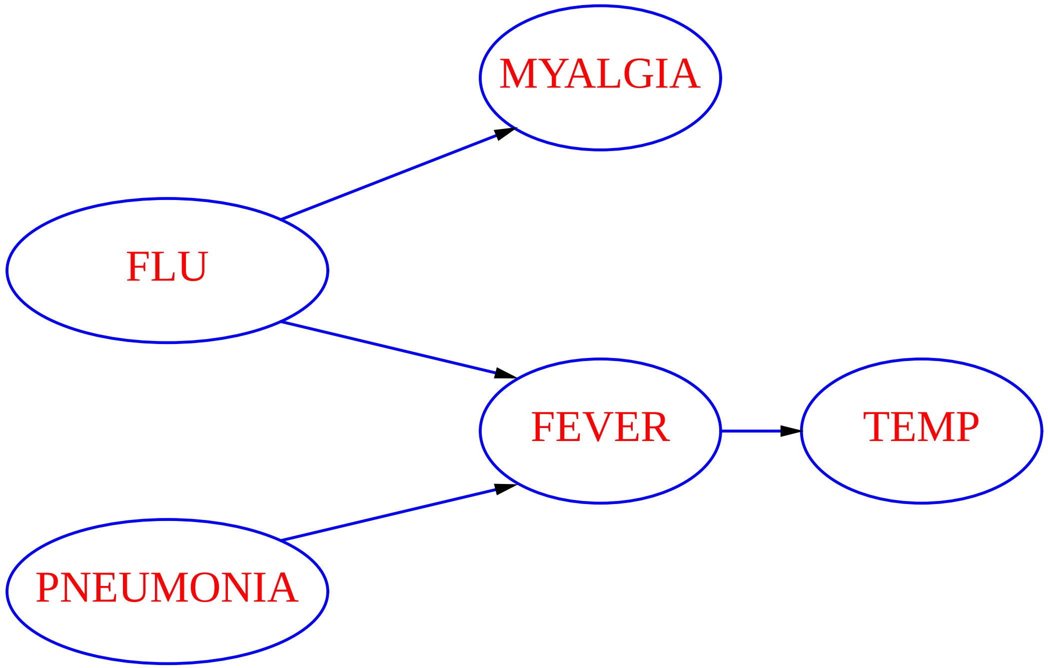
- A cause that has two or more outgoing arcs to other vertices is a common cause (factor) of those effects
- The effects of a common cause are usually observables (e.g. manifestations of failure of a device or symptoms in a disease)

Causal Graph: Example



- FEVER and PNEUMONIA are two alternative causes of fever (but may enhance each other)
- FLU has two common effects: MYALGIA and FEVER
- High body TEMPerature is an indirect effect of FLU and PNEUMONIA, caused by FEVER

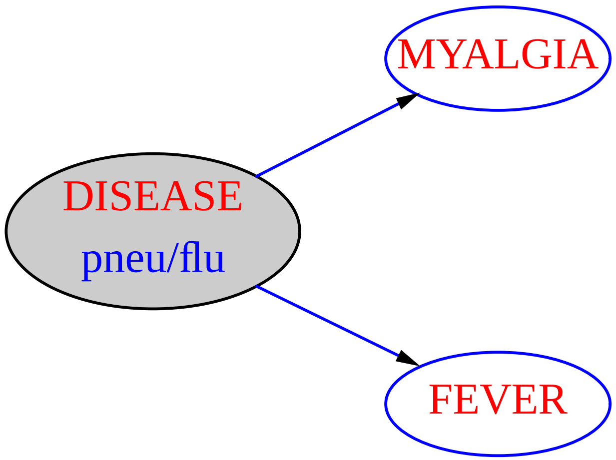
Check Independences



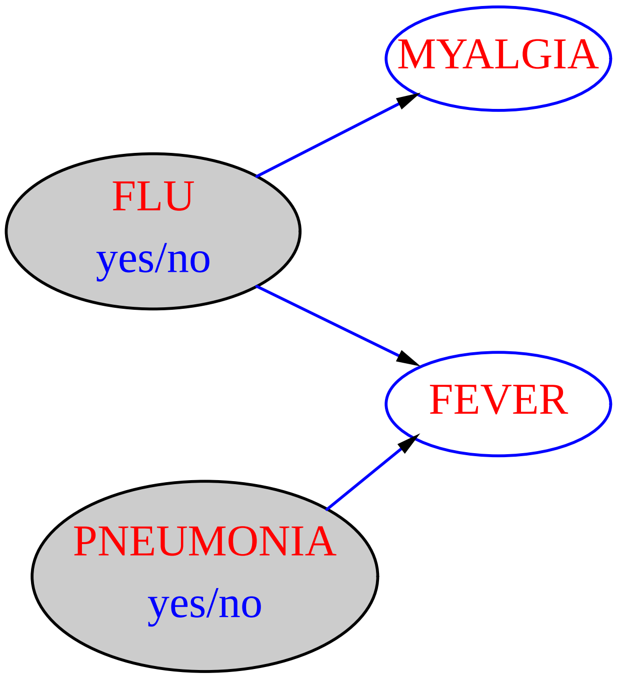
- Conditional independence: $X \perp\!\!\!\perp Y \mid Z$
  - $\text{FLU} \perp\!\!\!\perp \text{TEMP} \mid \text{FEVER}$
  - $\text{FEVER} \perp\!\!\!\perp \text{MYALGIA} \mid \text{FLU}$
  - $\text{PNEUMONIA} \perp\!\!\!\perp \text{FLU} \mid \varnothing$
  - $\text{PNEUMONIA} \not\!\perp\!\!\!\perp \text{FLU} \mid \text{FEVER}$

Choose Variables

- Factors are mutually exclusive (cannot occur together with absolute certainty): put as values in the same variable, or
- Factors may co-occur: multiple variables



(a) Single variable

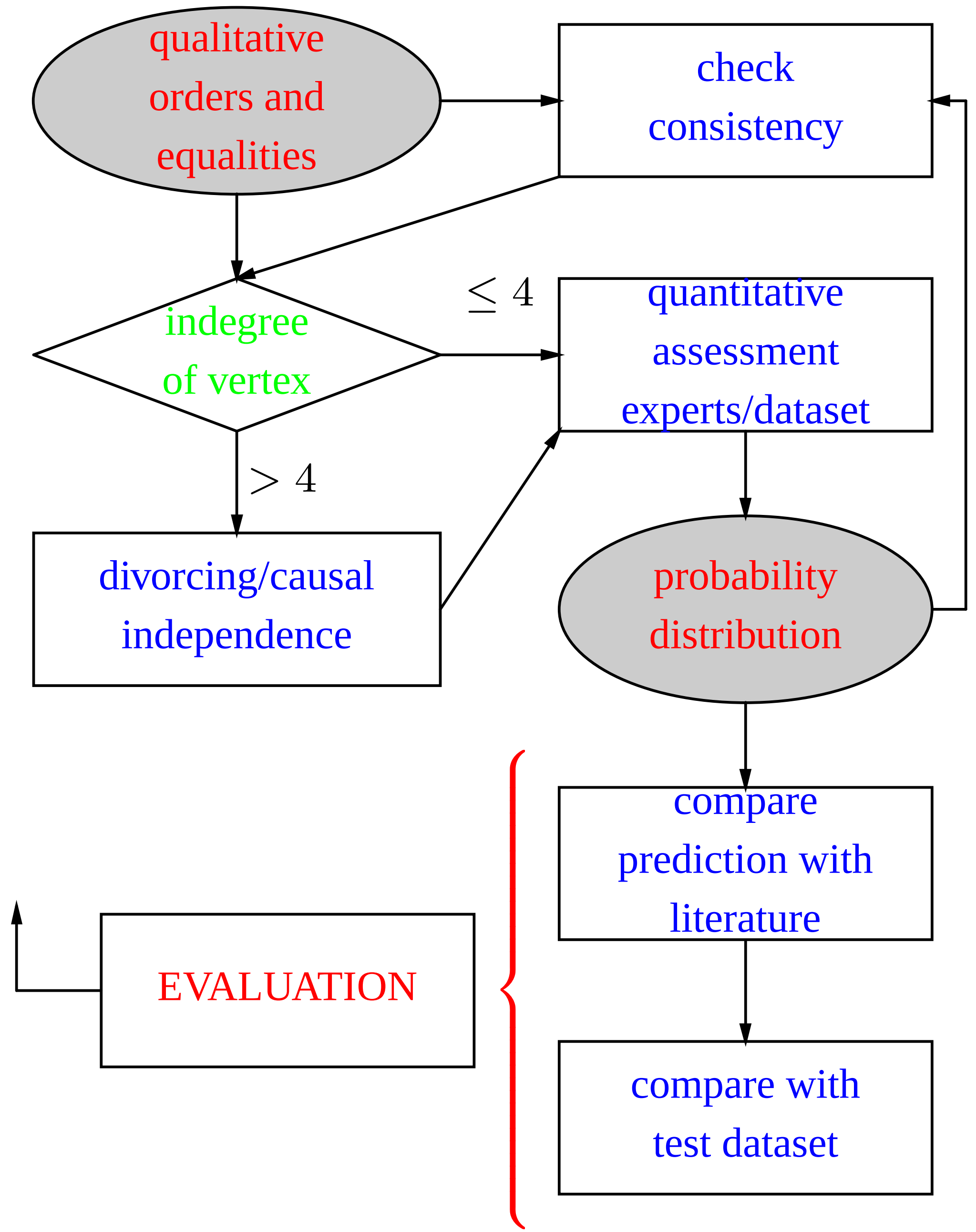


(b) Multiple variables

Choose Value Domains

- Discrete values
  - Mutually exclusive and exhaustive
  - Types:
    - binary, e.g., FLU $=$ *yes/no, true/false, 0/1*
    - ordinal, e.g., INCOME $=$ *low, medium, high*
    - nominal, e.g., COLOR $=$ *brown, green, red*
    - integral, e.g., AGE $= \{1, \ldots, 120\}$
- Continuous values
- Discretisation (of continuous and integral values)
  - Example for TEMP:
    $[-50, +5) \longrightarrow$ *cold*
    $[+5, +20) \longrightarrow$ *mild*
    $[+20, +50] \longrightarrow$ *hot*

Probability Assessment



Expert Judgements

- Qualitative probabilities:
  - *Qualitative orders*:

AGE	$P(\text{General Health Status} \mid \text{AGE})$
10-69	good $>$ average $>$ poor
70-79	average $>$ good $>$ poor
80-89	average $>$ poor $>$ good
≥ 90	poor $>$ average $>$ good

  - *Equalities*:

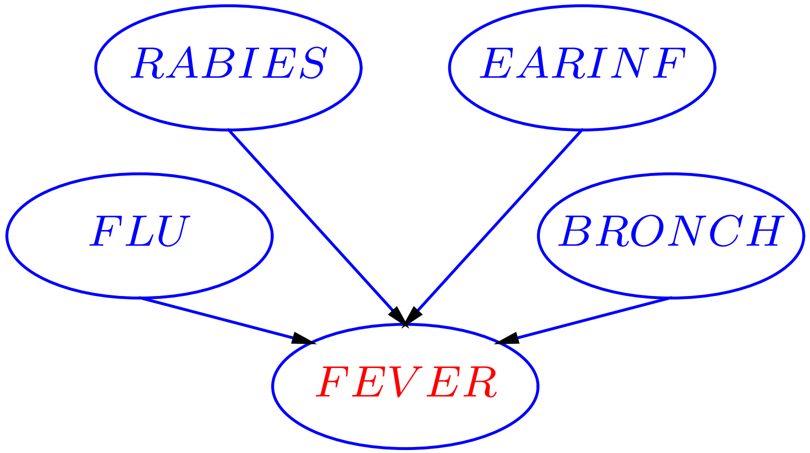
$$P(\text{CANCER} = T1 | \text{AGE} = 15 - 29) =$$
$$P(\text{CANCER} = T2 | \text{AGE} = 15 - 29)$$

Expert Judgements (cont.)

- Quantitative, subjective probabilities:

	$P(\text{GHS} \mid \text{AGE})$		
AGE	good	average	poor
10-69	0.99	0.008	0.002
70-79	0.3	0.5	0.2
80-89	0.1	0.5	0.4
≥ 90	0.1	0.3	0.6

A Bottleneck



- The number of parameters for the effect given n causes grows exponentially: 2^n for binary causes
- Unlikely evidence combination: $P(\text{FEVER}|\text{FLU}, \text{RABIES}, \text{EAR\_INFECTION}) = ?$

Problem: for many BNs too many probabilities have to be assessed

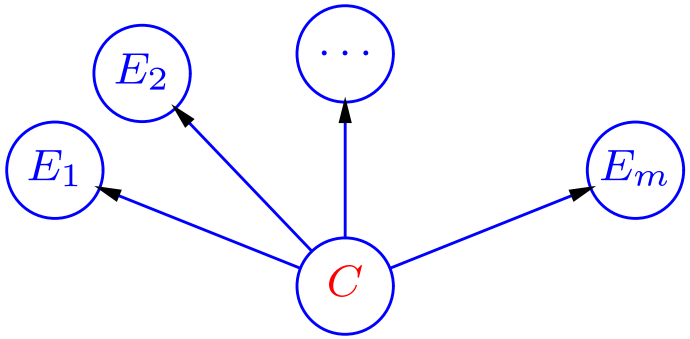
Special Form BN

Solution: use simpler probabilistic model, such that either

- the structure becomes simpler, e.g.,
  - *naive* (independent) form BN
  - *T*ree-*A*ugmented Bayesian *N*etwork (TAN)

  or,
- the assessment of the conditional probabilities becomes simpler (even though the structure is still complex), e.g.,
  - *parent divorcing*
  - *causal independence* BN

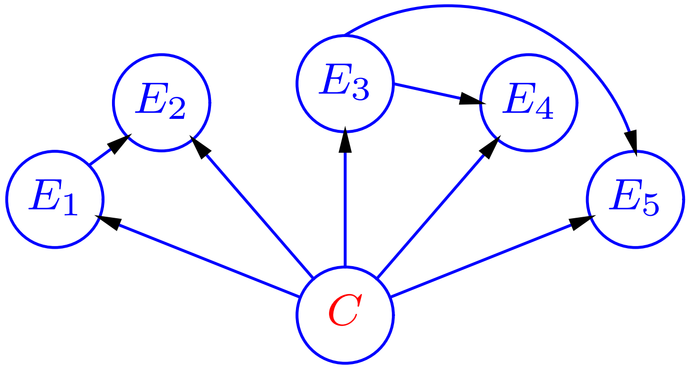
Independent (Naive) Form



- C is a class variable
- E_i are evidence variables and $\mathcal{E} \subseteq \{E_1, \ldots, E_m\}$. We have $E_i \perp\!\!\!\perp E_j \mid C$, for $i \neq j$. Hence, using Bayes' rule:

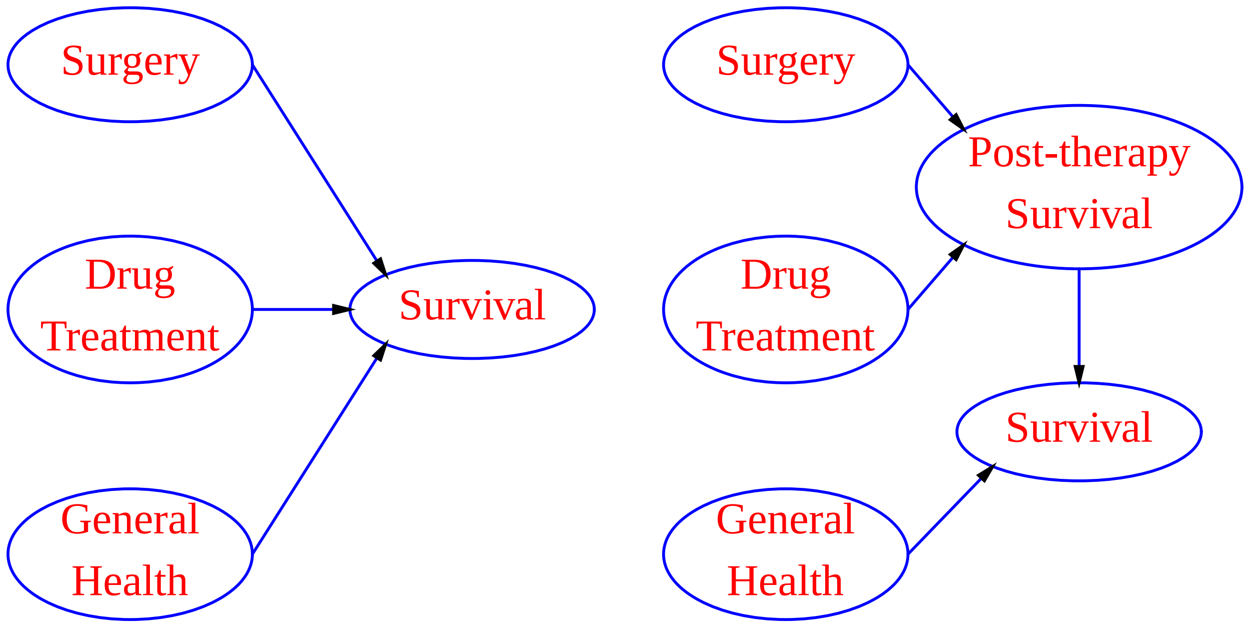
$$
\begin{aligned}
P(C \mid \mathcal{E}) &= \frac{P(\mathcal{E} \mid C) P(C)}{P(\mathcal{E})} && \text{with:} \\
P(\mathcal{E} \mid C) &= \prod_{E \in \mathcal{E}} P(E \mid C) && \text{by cond. ind.} \\
P(\mathcal{E}) &= \sum_{C} P(\mathcal{E} \mid C) P(C) && \text{marg. \& cond.}
\end{aligned}
$$

Tree-Augmented BN (TAN)



- Extension of Naive Bayes: reduce the number of independent assumptions
- Each node has at most two parents (one is the class node)

Divorcing Multiple Parents

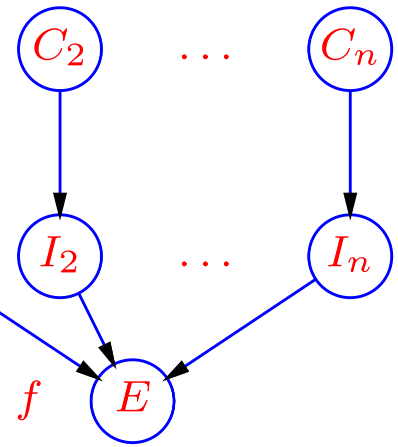


(a) Original network (b) Divorced network

Reduction in number of probabilities to assess:

- Identify a potential common effect of two or more parent vertices of a vertex
- Introduce a new variable into the network, representing the common effect

Causal Independence

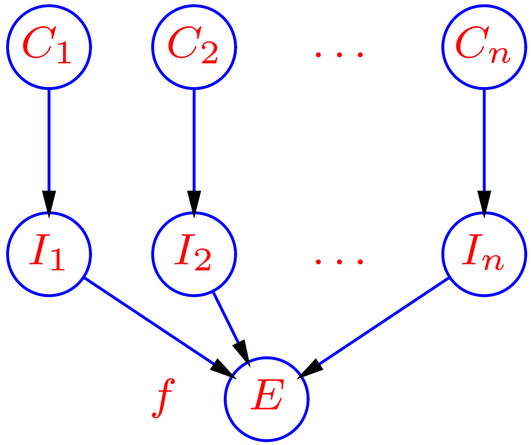


with:

- cause variables C_j, intermediate variables I_j, and the effect variable E
- $P(E \mid I_1, \ldots, I_n) \in \{0, 1\}$
- interaction function f, defined such that

$$f(I_1, \ldots, I_n) = \begin{cases} e & \text{if } P(e \mid I_1, \ldots, I_n) = 1 \\ \neg e & \text{if } P(e \mid I_1, \ldots, I_n) = 0 \end{cases}$$

Causal Independence

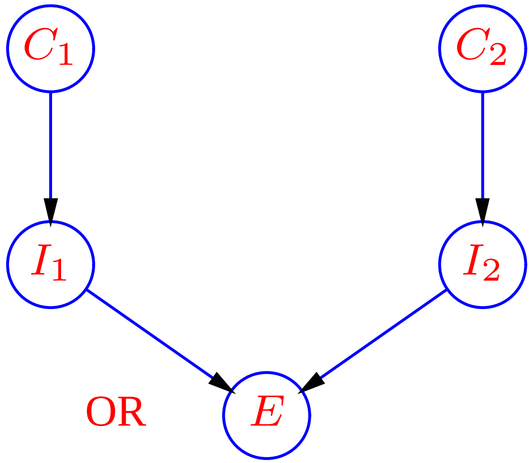


$$
\begin{aligned}
P(e \mid C_1, \ldots, C_n) &= \sum_{I_1,\ldots,I_n} P(e \mid I_1, \ldots, I_n) P(I_1, \ldots, I_n \mid C_1, \ldots, C_n) \\
&= \sum_{f(I_1,\ldots,I_n)=e} P(e \mid I_1, \ldots, I_n) P(I_1, \ldots, I_n \mid C_1, \ldots, C_n)
\end{aligned}
$$

Note that as $I_i \perp\!\!\!\perp I_j \mid \varnothing$, and $I_i \perp\!\!\!\perp C_j \mid C_i$, for $i \neq j$, it holds that:

$$
P(I_1, \ldots, I_n \mid C_1, \ldots, C_n) = \prod_{k=1}^{n} P(I_k \mid C_k)
$$

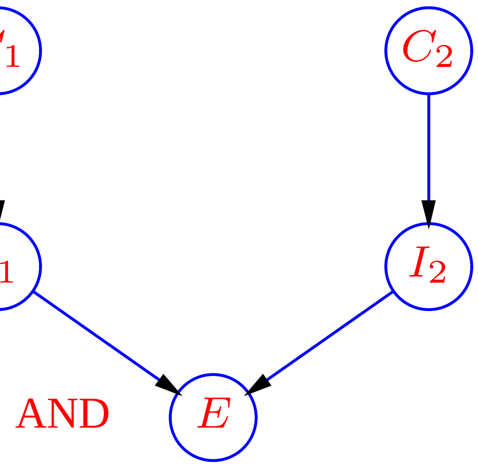
Noisy OR



- Interactions among causes, as represented by the function f and $P(E \mid I_1, I_2)$, is a logical OR
- Meaning: presence of any one of the causes C_i with absolute certainty will cause the effect e (i.e., $E = true$)

$$\begin{aligned} P(e|C_1, C_2) &= \sum_{I_1 \vee I_2 = e} P(e|I_1, I_2) \prod_{k=1,2} P(I_k|C_k) \\ &= P(i_1|C_1)P(i_2|C_2) + P(\neg i_1|C_1)P(i_2|C_2) \\ &\quad + P(i_1|C_1)P(\neg i_2|C_2) \end{aligned}$$

Noisy AND



- Interactions among causes, as represented by the function f and $P(E \mid I_1, I_2)$, is a logical AND
- Meaning: presence of all causes C_i with absolute certainty will cause the effect e (i.e. $E = true$); otherwise, $\neg e$

$$P(e|C_1, C_2) = \cdots$$

Model Refinement

Model refinement is necessary:

- How:
  - Manual
  - Automatic: sensitivity analysis
- What:
  - Probability adjustment
  - Removing irrelevant factors
  - Adding previously hidden, unknown factors
  - Causal relationships adjustment, e.g., including, removing independence relations