Bayesian Network Modelling and Clinican Decision Making in Liver Disease

Peter Lucas

peterl@cs.ru.nl

Department of Model-based System Development

Institute for Computing and Information Sciences

Radboud University Nijmegen, the Netherlands

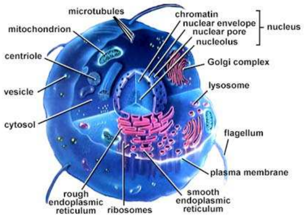
Background

Probabilistic graphical models, such as Bayesian networks, can be used for:

- systems modelling and simulation
- knowledge discovery (learning)
- least commitment principle

Integration:

- molecular, (sub)cellular biology
- patient, environment levels



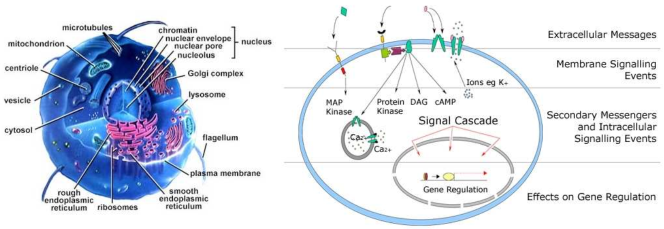
Uncertainty:

- individual variation

Challenges

Translational medicine:

- To link basic scientific discoveries to clinical research
- To translate results from clinical research to clinical practice
- Clinical practice:
  - Diagnosis
  - Treatment, prognosis
  - Follow-up/monitoring



Diagnosis of liver disease

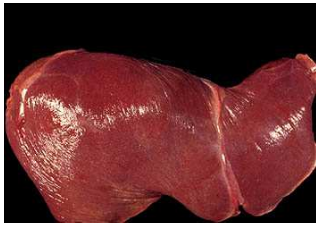
Clinical point of view:

(1) The disorder is primarily affecting the hepatocytes (hepatocellular disorder) or the biliary tract (biliary obstructive disorder)

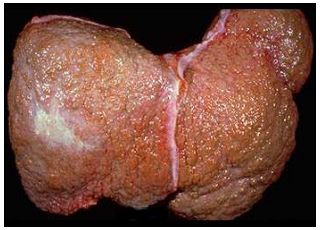
(2) disorder is acute or chronic in nature

(3) disorder has benign or malignant features

Based in this: plan for further diagnostic assessment



acute (hepatitis)



chronic (cirrhosis)

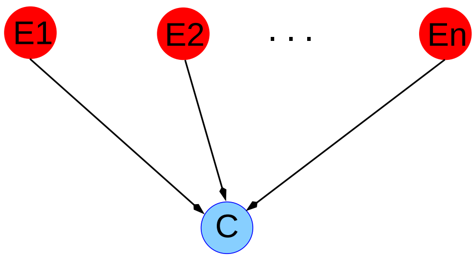
malignant

Pocket diagnostic chart

	No vs. Ob	Ac vs. Ch	Be vs. Ma
Age: 31 – 64 years	+7	+5	
$\geq$ 65 years	+12	+5	
Previous history:			
Jaundice due to cirrhosis	−7	+8	
Cancer in GI-tract, pancreas, bile system, or breast	+10		+7
Leukaemia or malignant lymphoma	−13		
Previous biliary colics or proven gallstones	+3	+7	−7
In treatment for congestive heart failure		−5	
Present history:			
$\geq$ 2 weeks			+7
Upper abdominal pain:			
sever	+9		−6
slight or moderate	+4		
Fever:			
without chills		−3	−5
with chills		−6	−10
Intermittent jaundice	+5		−5
Weight loss ($\geq$ 2 kg)			+4
Alcohol:			
1 – 4 drinks per day	−4		
$\geq$ 5 drinks per day	−4	+4	
SUM left			

	No vs. Ob	Ac vs. Ch	Be vs. Ma
Physical examination:			
Spiders	−6	+11	
Ascites	−3	+6	
Liver surface nodular		+5	
Gall bladder:			
Courvoisier	+16		+11
firm or tender	+5		
Clinical chemistry:			
bilirubin $\geq 200\mu$mol/l	+5	−5	+5
Alkaline phosphatase:			
400 – 1000 U/l	+6		
> 1000 U/l	+11		+6
ASAT:			
40 – 319 U/l		+5	
$\geq$ 320 U/l	−10	+1	+6
Clotting factors:			
$\leq$ 0.55		+8	+5
0.56 – 0.70		+5	+5
LDH $\geq$ 1300 U/l		−5	+7
SUM left			
CONSTANTS	−19	−21	−8
TOTAL SCORE			

- P. Matzen, et al. Liver 4 (1984) 360–71
- Accuracy: 75–77% of patients with jaundice
- Logistic regression: $S_c = \sum_k^{n_c} \omega_k^c e_k^c$, $c =$ non-obstructive, acute, benign, with $P(c \mid \mathcal{E}) = [1 + \exp -S_c]^{-1}$
- As Bayesian network: $P(C, E_1, \ldots, E_n) = P(C|E_1, \ldots, E_n) \times P(E_1, \ldots, E_n)$



Requirements modelling language

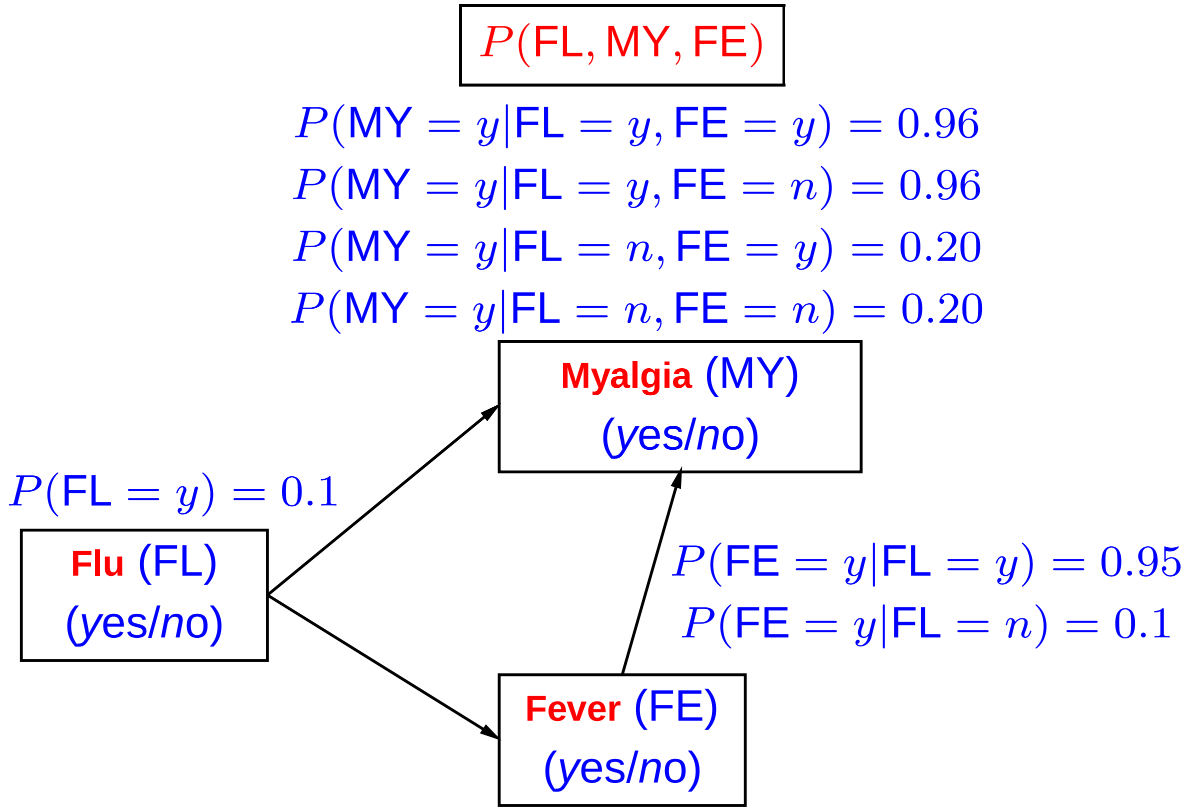
Language for disease modelling should include:

- Variables X, Y
- Interactions among variables $(X_1, \ldots, X_n) \to Y$
- Possibility to attach meaning to interactions in terms of causality
- Allow coping with uncertainty

$\Rightarrow$ Probabilistic graphical models

- Represent joint probabilility distribution $P(X_1, \ldots, X_n, Y)$
- Graphical representation: Markov models, Bayesian networks, chain graphs, $\ldots$

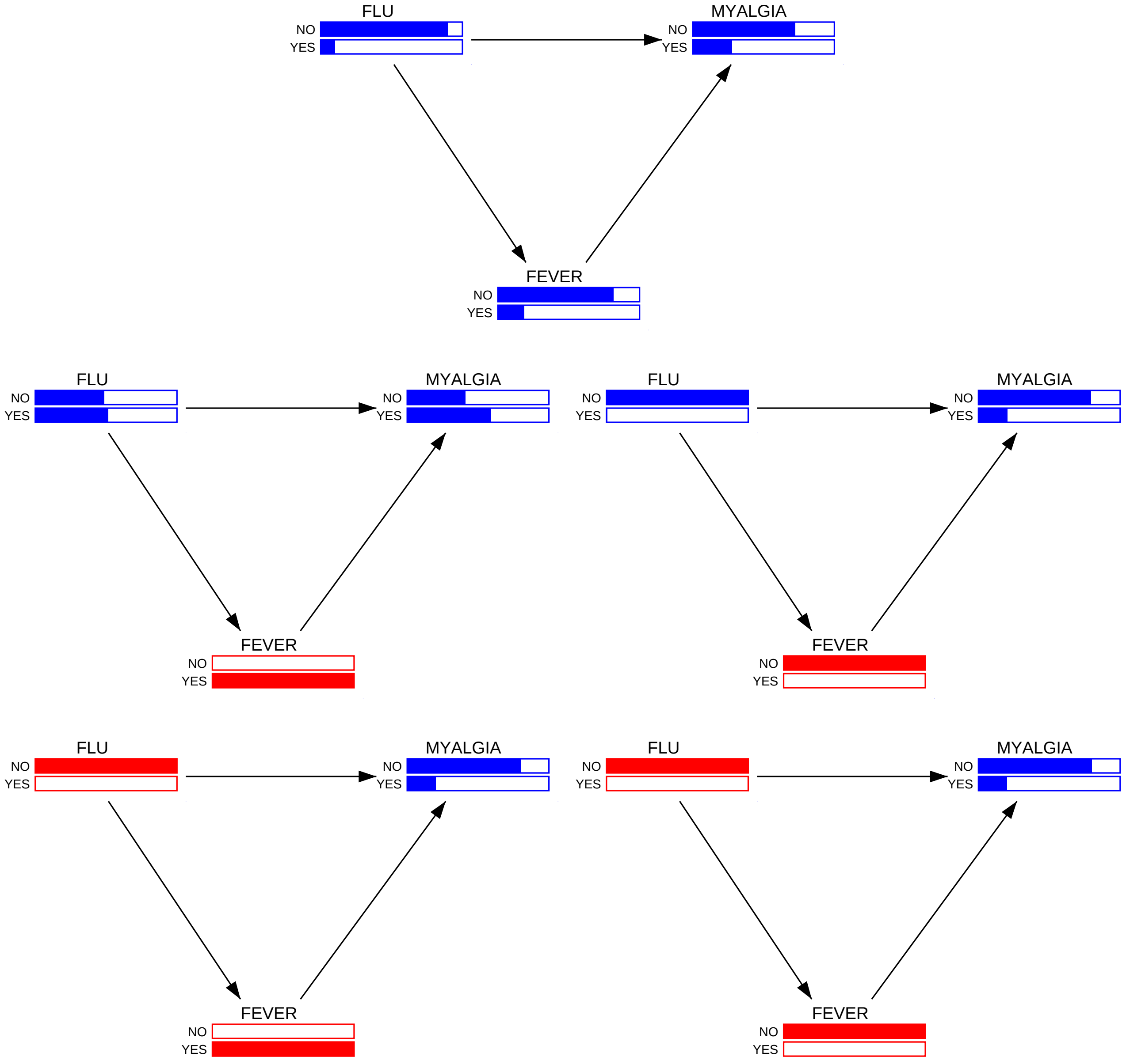
Bayesian network



$$\begin{aligned} P(\text{FL}, \text{MY}, \text{FE}) &= P(\text{MY} \mid \text{FL}, \text{FE}) P(\text{FE} \mid \text{FL}) P(\text{FL}) \\ &= P(\text{MY} \mid \text{pa}(\text{MY})) P(\text{FE} \mid \text{pa}(\text{FE})) P(\text{FL} \mid \text{pa}(\text{FL})) \end{aligned}$$

Example: $P(\neg \textit{fl}, \textit{my}, \textit{fe}) = 0.20 \cdot 0.1 \cdot 0.9 = 0.018$

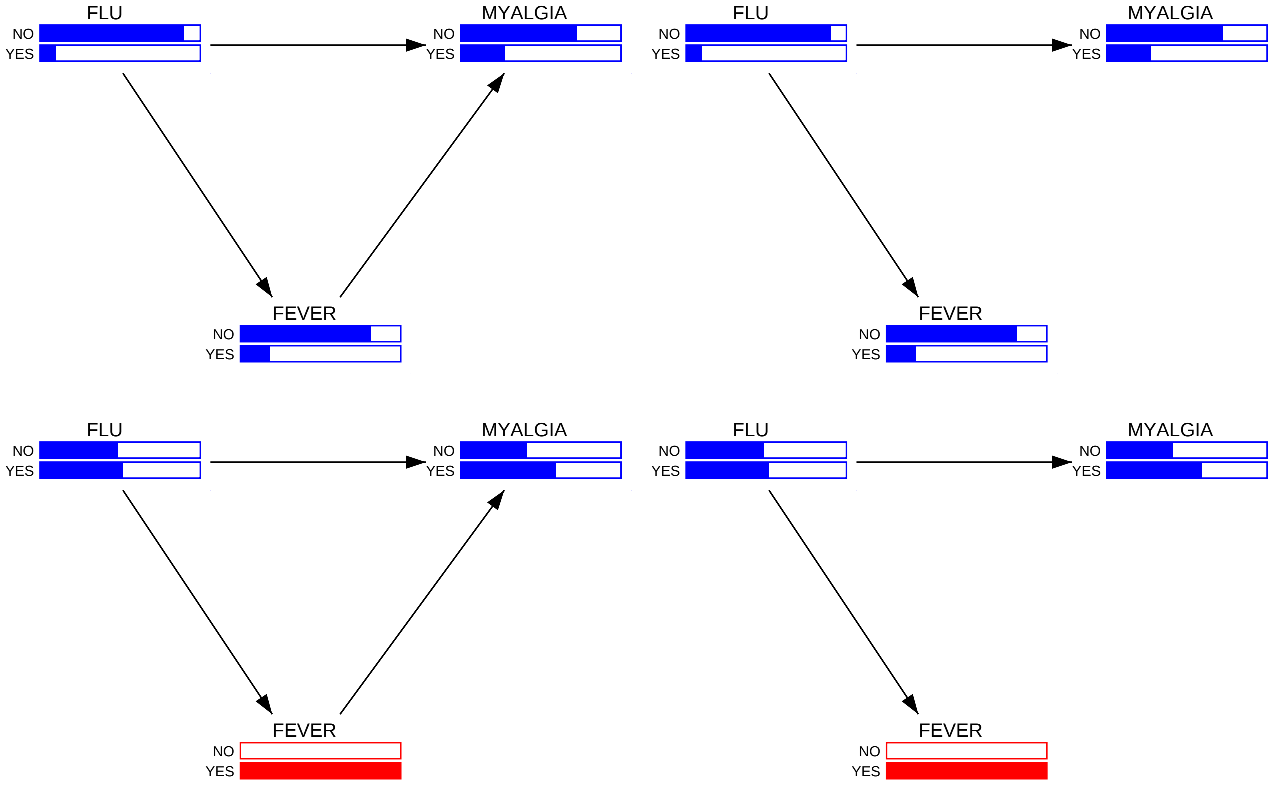
Independence and reasoning



Independence and reasoning

Arc from FEVER to MYALGIA can be removed, hence

$$P(\mathsf{MY} \mid \mathsf{FL}) \; (= P(\mathsf{MY} \mid \mathsf{FL}, \mathsf{FE}))$$



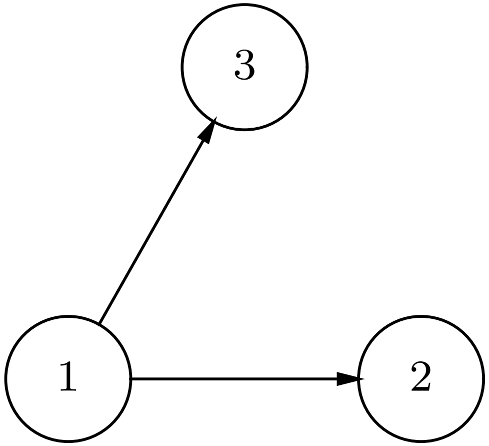
Independence relation

Let P be a probability distribution of X then U is called conditionally independent of Y given Z, denoted as

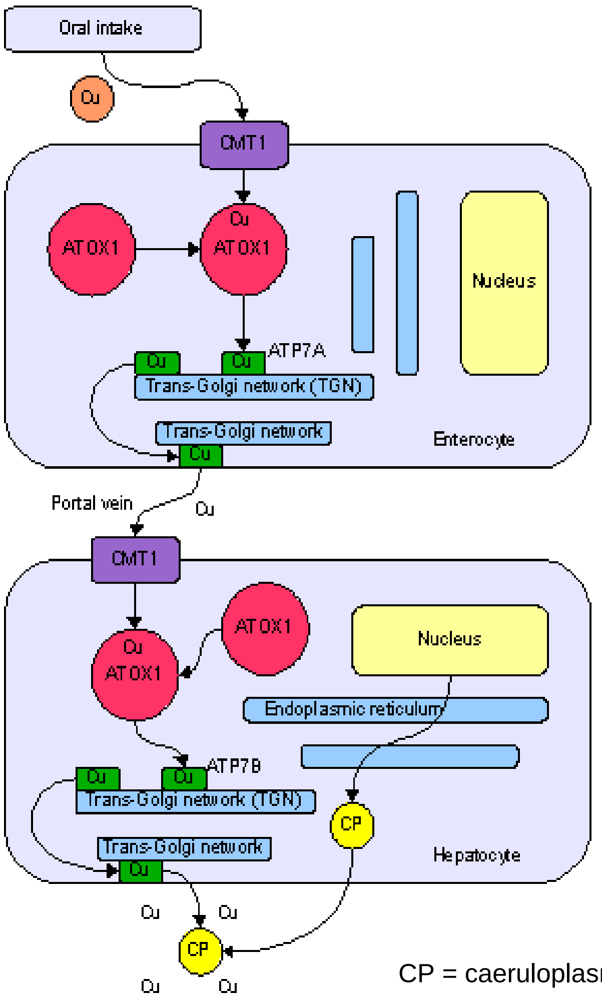
$$U \perp\!\!\!\perp Y \mid Z, \quad \text{iff} \quad P(U \mid Y, Z) = P(U \mid Z)$$

Note: This relation is completely defined in terms of the probability distribution P, but there is a relationship to graphs, for example:

$$X_2 \perp\!\!\!\perp X_3 \mid X_1$$

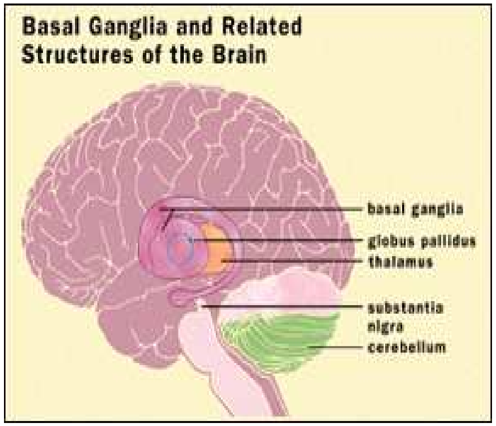


Wilson’s disease



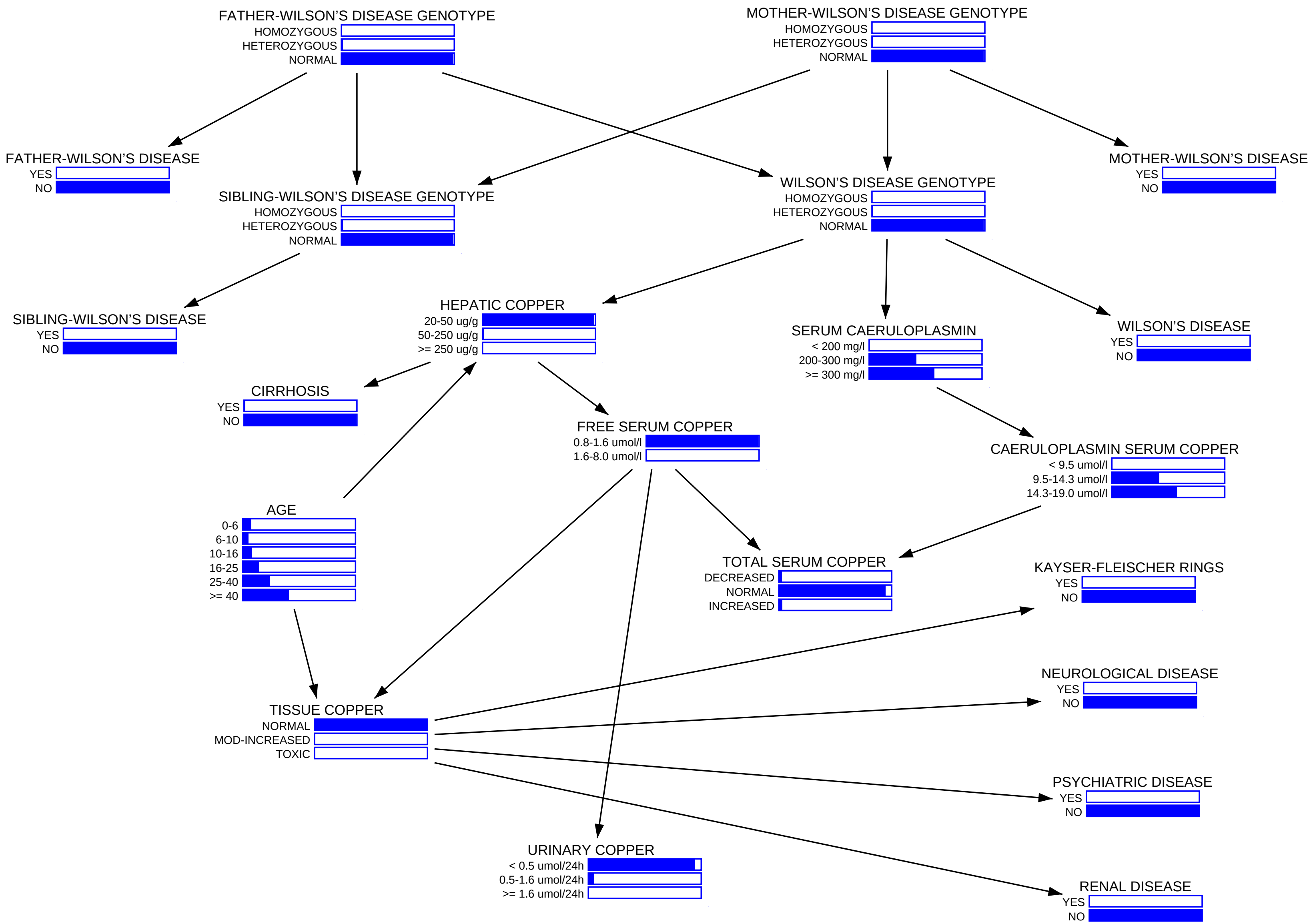
CP = caeruloplasmin

Cearuloplasmin

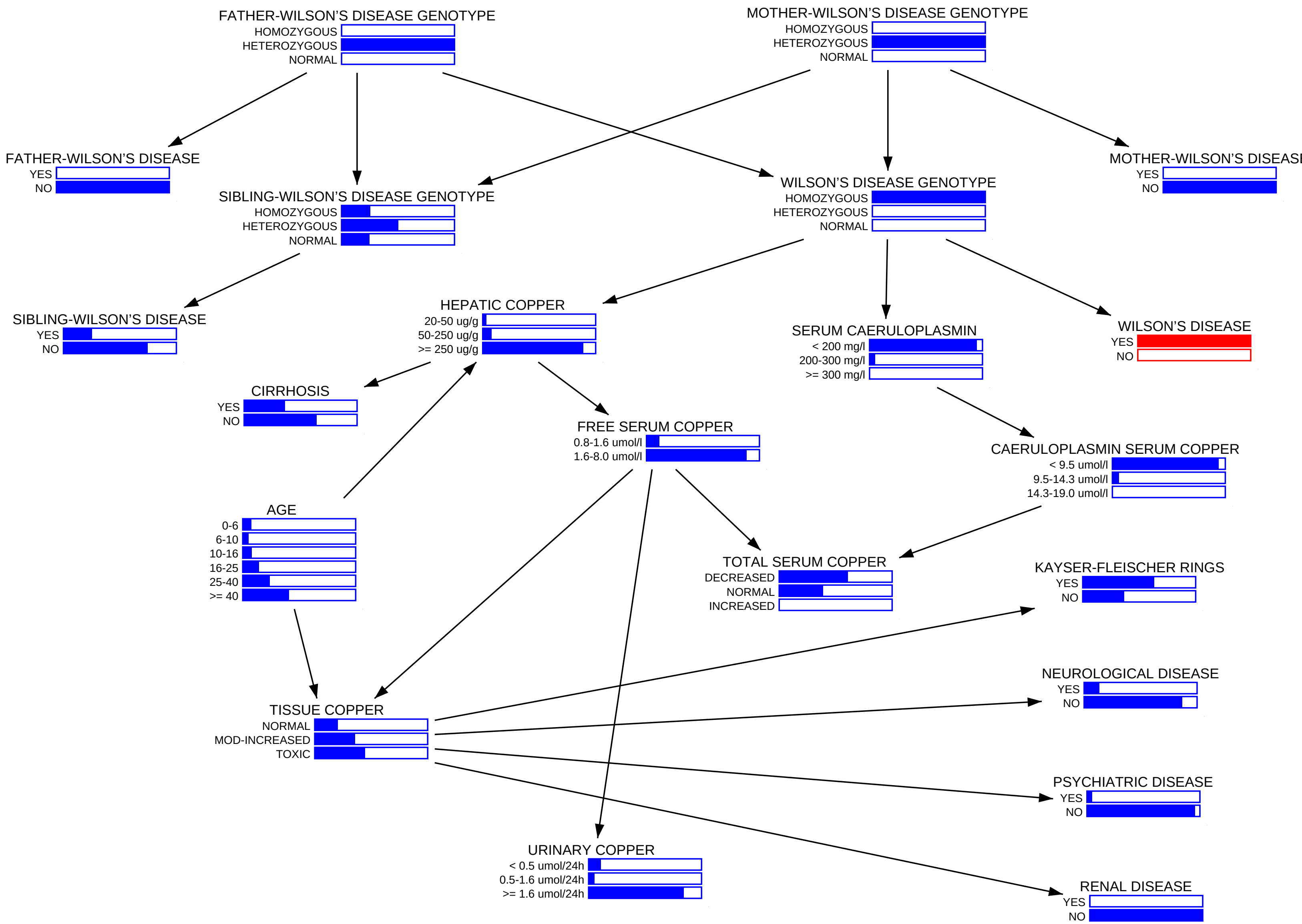


Kayser−Fleischer rings

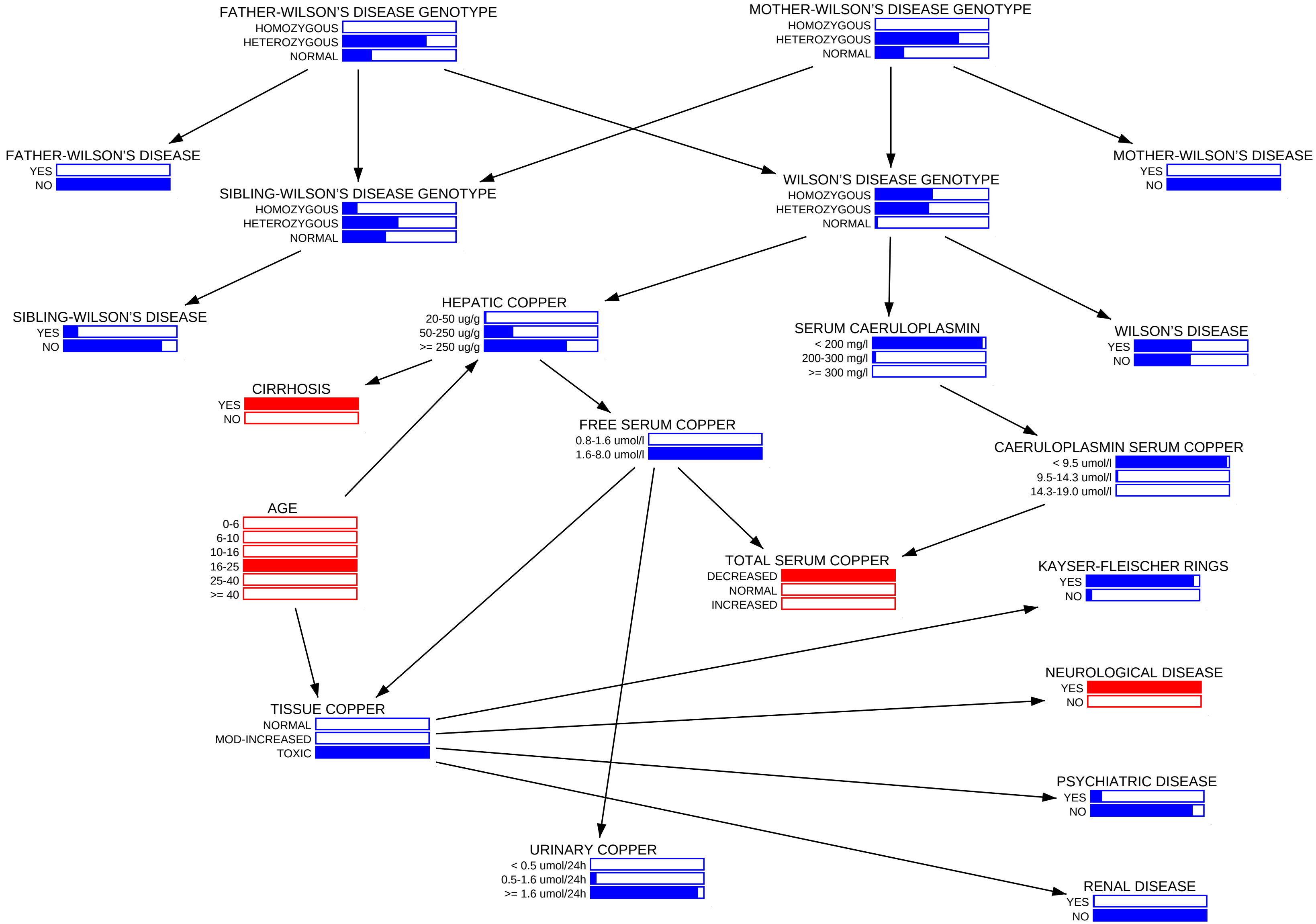
Wilson's disease network – prior



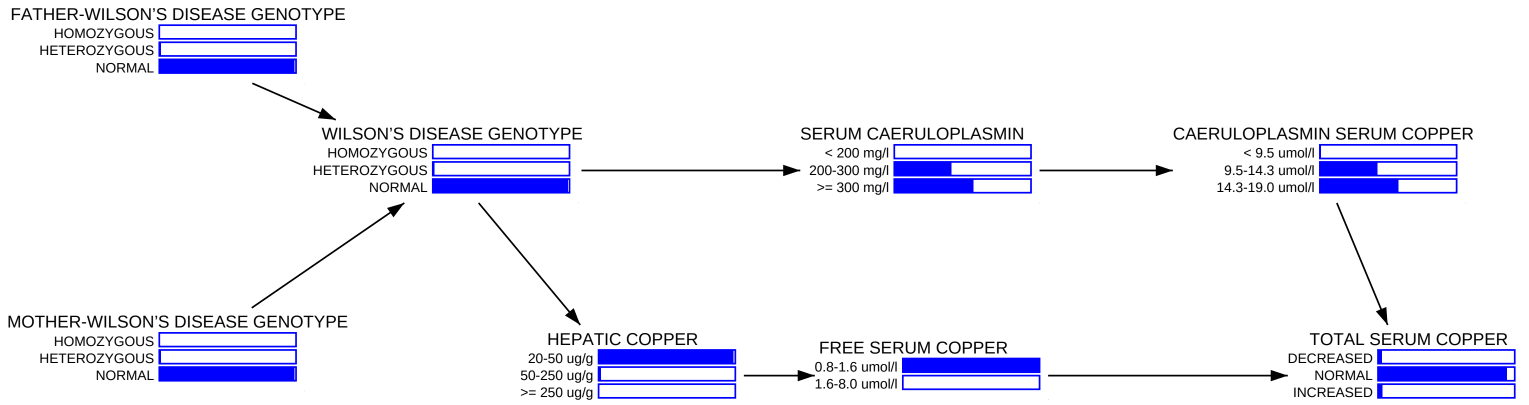
Network prediction



Network posterior



Reading off the independences

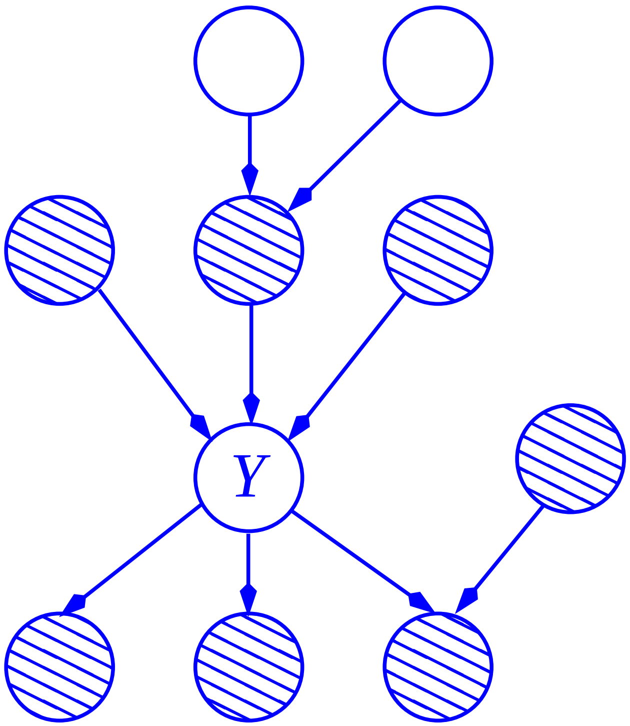


Examples:

- $\text{FWDG} \perp\!\!\!\perp \text{MWDG} \mid \varnothing$
- $\text{FWDG} \not\perp\!\!\!\perp \text{MWDG} \mid \text{WDG}$
- also: $\text{FWDG} \not\perp\!\!\!\perp \text{MWDG} \mid \text{HC}$
- $\text{WDG} \perp\!\!\!\perp \text{TSC} \mid \{\text{SC}, \text{FSC}\}$

(FWDG = Father Wilson's Disease Genotype, etc.)

Markov blanket

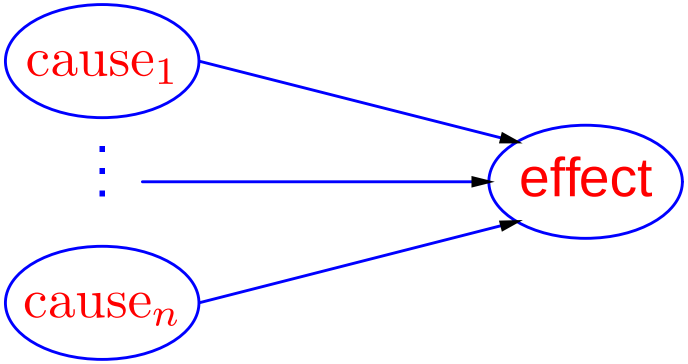


MB = Markov blanket: marked nodes

$$Y \perp\!\!\!\perp X \backslash (\{Y\} \cup \text{MB}(Y)) \mid \text{MB}(Y)$$

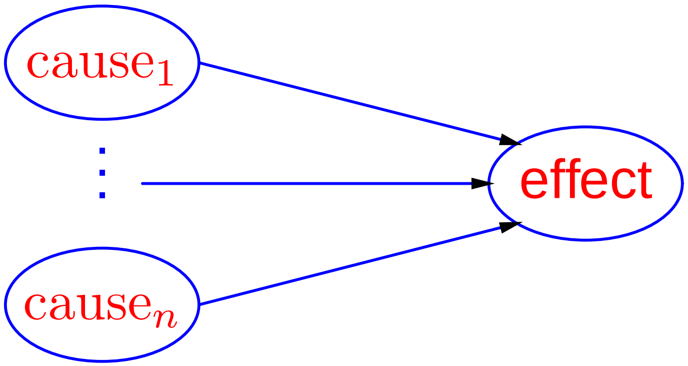
- The Markov blanket shields Y from all other factors, i.e. Markov blanket includes all factors that directly affect Y
- Has biological meaning

Causal graph: topology



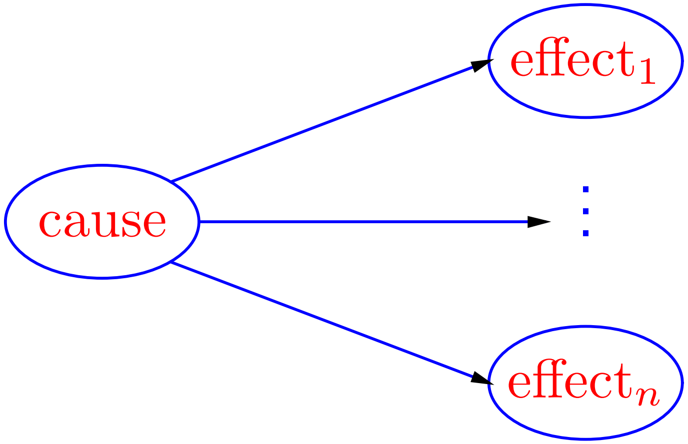
- Identify factors that are relevant
- Determine how those factors are causally related to each other
- The arrow 'cause → effect' does mean that 'cause' is a factor involved in causing 'effect'

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:
  - Positive synergy: Polution $\longrightarrow$ Cancer $\longleftarrow$ Smoking
  - Negative synergy: Vaccine $\longrightarrow$ Death $\longleftarrow$ Smallpox

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. signs and symptoms in a disease)

Specification of Interactions

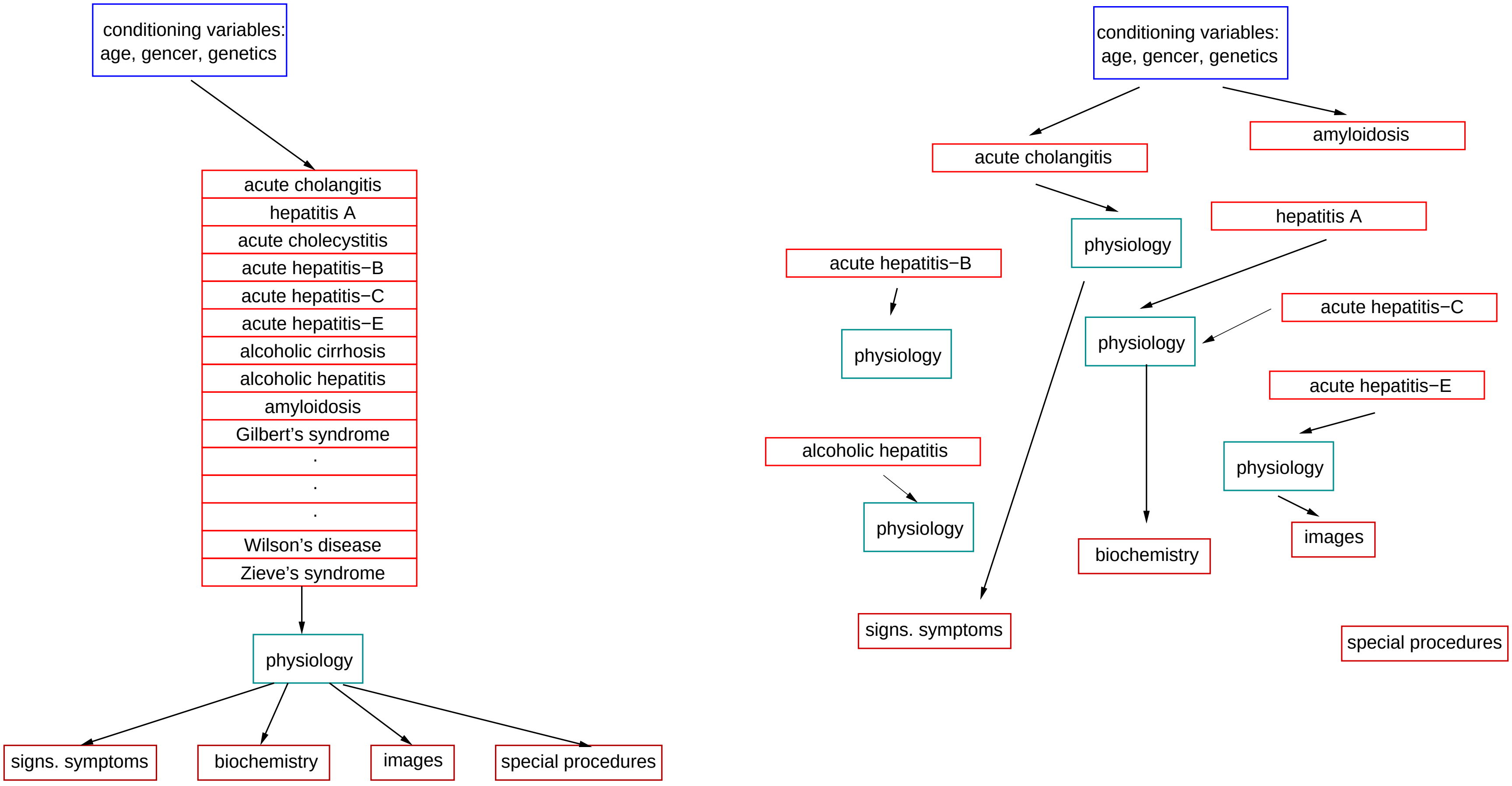
- Compact specification: probability tables

$$P(X_i \mid \text{parents}(X_i))$$

  can still be large even when taking into account independence information
- Easy way to map domain knowledge to entries into a probability table
- Way to use qualitative knowledge about interactions as constraints to probabilistic information
- Various techniques available to reduce size of specification

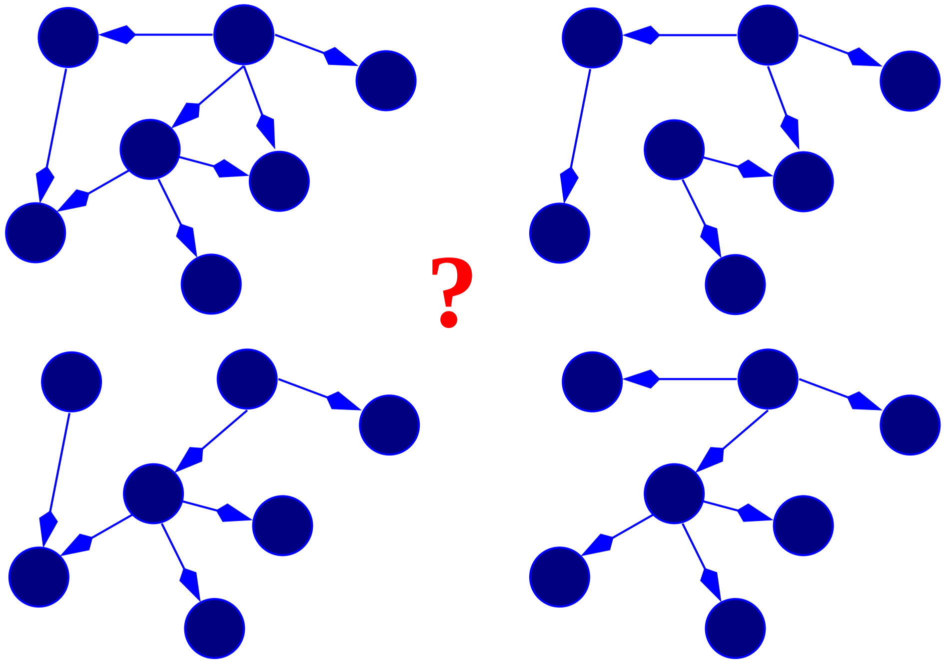
Diagnostic models (of liver disease)

Diagnosis: $d^* = \max_d P(d \mid \textit{Evidence})$ (for any disease)



$P(\text{acute hepatitis-B}, \text{Wilson's disease}) = 0$ $\quad$ $P(\text{acute hepatitis-B}, \text{Wilson's disease}) > 0$

Learning Bayesian networks



- Bayesian networks $\Leftrightarrow$ datasets?
- Learning:
  - parameter (distribution given structure) learning
  - structure (topology) learning

Comparing models

Let D be data, G be the structure and θ_G be the parameters of a BN; common methods:

- Likelihood: $L_{\theta_G}(G) = \Pr(D \mid G, \theta_G)$, for given G and θ_G. Estimating parameters by maximum log-likelihood: $l(G) = \max_{\theta_G} \log \Pr(D \mid G, \theta_G)$
- Marginal likelihood:

$$M(G) = \Pr(D \mid G) = \int_{\theta_G} \Pr(D \mid G, \theta_G) \Pr(\theta_G) \, \mathrm{d}\theta_G$$

with prior $\Pr(\theta_G)$ and parameters θ_G marginalised out

($\Pr$ is a density on data, structure, and parameters)

Conclusions

- PGMs: powerful for modelling for biomedicine:
  - white-box representation of interactions
  - can be learnt from data (structure and parameters)
  - handling of uncertainty in relationship
- Graph-based independence reasoning supplements probabilistic reasoning
- Very intuitive, software available (e.g. in R), and anyone can use PGMs after some training