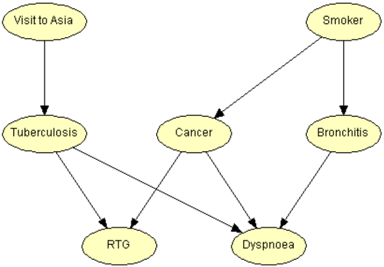
Probabilistic graphical models: current research activities

Jirka Vomlel

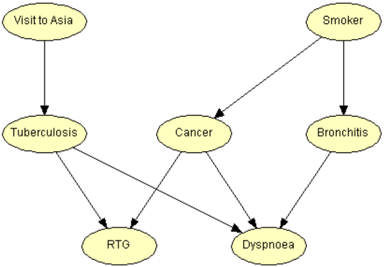
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
http://www.utia.cz/vomlel

Aalborg, Denmark, November, 20, 2013

A simple Bayesian network model - Chest Clinic



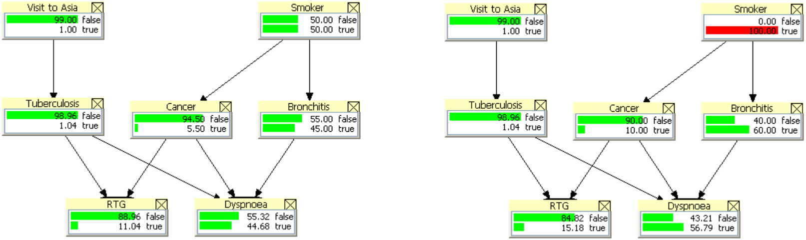
A simple Bayesian network model - Chest Clinic



Conditional probability tables (CPTs)

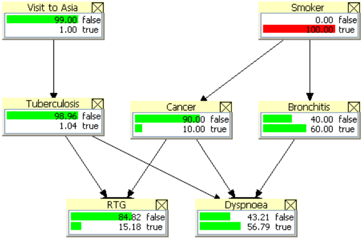
P(Visit to Asia)
P(Smoker)
P(Tuberculosis | Visit to Asia)
P(Cancer | Smoker)
P(Bronchitis | Smoker)
P(RTG | Tuberculosis, Cancer)
P(Dyspnoea | Tuberculosis, Cancer, Bronchitis)

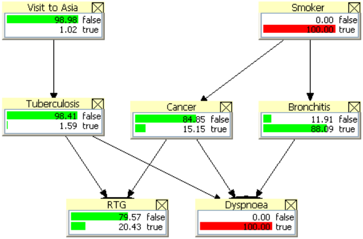
Probabilistic inference with the Bayesian network



$P(X|\text{Smoker=true})$

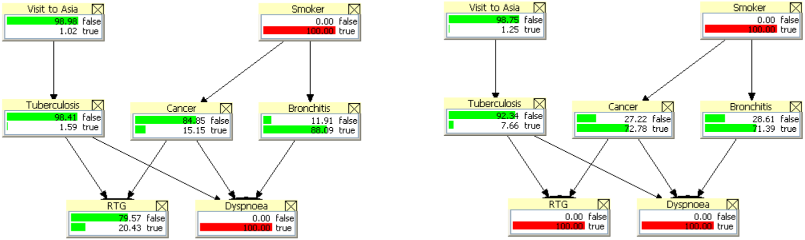
Probabilistic inference with the Bayesian network





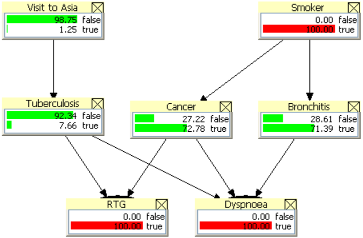
$P(X|$Smoker=true, Dyspnoea=true$)$

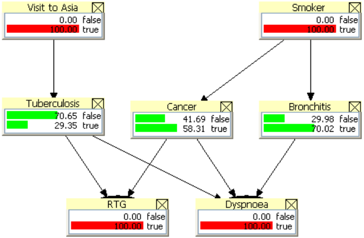
Probabilistic inference with the Bayesian network



$P(X|$Smoker=true, Dyspnoea=true, RTG=true$)$

Probabilistic inference with the Bayesian network





$P(X|$Smoker=true, Dyspnoea=true, RTG=true, Visit to Asia=true$)$

CPT $P(\text{RTG} \mid \text{Tuberculosis, Cancer})$

First, assume a deterministic function. RTG is positive iff the patient has tuberculosis or cancer.

RTG	Tuberculosis	Cancer	p
0	0	0	1
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	1

CPT $P(\text{RTG} \mid \text{Tuberculosis, Cancer})$

RTG can have other reasons for being positive and RTG need not be positive even if the patient has tuberculosis or cancer.

RTG	Tuberculosis	Cancer	p	p'	
0	0	0	1	p_0	0.95
0	0	1	0	$p_0 * p_1$	0.019
0	1	0	0	$p_0 * p_2$	0.019
0	1	1	0	$p_0 * p_1 * p_2$	0.00038
1	0	0	0	$1 - p_0$	0.05
1	0	1	1	$1 - p_0 * p_1$	0.981
1	1	0	1	$1 - p_0 * p_2$	0.981
1	1	1	1	$1 - p_0 * p_1 * p_2$	0.99962
				$p_0, p_1, p_2 \in \langle 0, 1 \rangle$	

CPT P(RTG | Tuberculosis, Cancer)

RTG can have other reasons for being positive and RTG need not be positive even if the patient has tuberculosis or cancer.

RTG	Tuberculosis	Cancer	p	p'	
0	0	0	1	p_0	0.95
0	0	1	0	$p_0 * p_1$	0.019
0	1	0	0	$p_0 * p_2$	0.019
0	1	1	0	$p_0 * p_1 * p_2$	0.00038
1	0	0	0	$1 - p_0$	0.05
1	0	1	1	$1 - p_0 * p_1$	0.981
1	1	0	1	$1 - p_0 * p_2$	0.981
1	1	1	1	$1 - p_0 * p_1 * p_2$	0.99962
				$p_0, p_1, p_2 \in \langle 0, 1 \rangle$	

This local model is called "noisy-or".

CPT $P(\text{RTG} \mid \text{Tuberculosis, Cancer})$

RTG can have other reasons for being positive and RTG need not be positive even if the patient has tuberculosis or cancer.

RTG	Tuberculosis	Cancer	p	p'	
0	0	0	1	p_0	0.95
0	0	1	0	$p_0 * p_1$	0.019
0	1	0	0	$p_0 * p_2$	0.019
0	1	1	0	$p_0 * p_1 * p_2$	0.00038
1	0	0	0	$1 - p_0$	0.05
1	0	1	1	$1 - p_0 * p_1$	0.981
1	1	0	1	$1 - p_0 * p_2$	0.981
1	1	1	1	$1 - p_0 * p_1 * p_2$	0.99962
				$p_0, p_1, p_2 \in \langle 0, 1 \rangle$	

This local model is called "noisy-or".
Let k be the number of parents. We need to specify $k + 1$ values $p_0, p_1, \ldots, p_k$ instead of 2^k in a general CPT.

Current research activities

- Model elicitation

Current research activities

- Model elicitation
  - learning models from data (using Integer Programming)
  - learning models with local structure of a noisy-or like type.
  - combination of expert knowledge and data (biological pathways and experimental data)

Current research activities

- Model elicitation
- Efficient inference with special types of probabilistic models

Current research activities

- Model elicitation
- Efficient inference with special types of probabilistic models
  - exploiting determinism
  - exploiting local structure of CPTs

Current research activities

- Model elicitation
- Efficient inference with special types of probabilistic models
- Methods of approximate inference

Current research activities

- Model elicitation
- Efficient inference with special types of probabilistic models
- Methods of approximate inference
  - iterative refinement
  - anytime inference methods

Current research activities

- Model elicitation
- Efficient inference with special types of probabilistic models
- Methods of approximate inference
- Other types of probabilistic graphical models:

Current research activities

- Model elicitation
- Efficient inference with special types of probabilistic models
- Methods of approximate inference
- Other types of probabilistic graphical models:
  – models with continuous variables (other than Gaussian)
  – models with causal interpretation of directed edges
  – models with both directed and undirected edges in the model (e.g. chain graphs)
  – modeling temporal and spatial information.

Current research activities

- Model elicitation
- Efficient inference with special types of probabilistic models
- Methods of approximate inference
- Other types of probabilistic graphical models:
- Finding good strategies with the help of a BN:

Current research activities

- Model elicitation
- Efficient inference with special types of probabilistic models
- Methods of approximate inference
- Other types of probabilistic graphical models:
- Finding good strategies with the help of a BN:
  - Decision-Theoretic Troubleshooting
  - Adaptive Testing

Current research activities

- Model elicitation
- Efficient inference with special types of probabilistic models
- Methods of approximate inference
- Other types of probabilistic graphical models:
- Finding good strategies with the help of a BN:
- Classification and regression for medical applications:

Current research activities

- Model elicitation
- Efficient inference with special types of probabilistic models
- Methods of approximate inference
- Other types of probabilistic graphical models:
- Finding good strategies with the help of a BN:
- Classification and regression for medical applications:
  - mortality prediction
  - prediction of medical care costs