UP2DATE – January 30th, 2009

Bayesian Networks

Jan Lemeire

Department of Electronics and Informatics

Vrije Universiteit Brussel

jan.lemeire@vub.ac.be

www.vub.ac.be

Up2Date technology

Overview

- Introduction: From Bayes' rule to Bayesian networks
- The use of Bayesian networks
- Learning from data
- Comparison with other techniques
- Conclusions

Belief Update during Medical Diagnosis

Information	P(cancer)
Prior belief	0.001
Short of breath (symptom)	0.002
Smoker	0.008
X-rays (clinical test)	0.04
It is not bronchitis	0.5

www.mastersininnovation.com

MASTERS IN INNOVATION

Bayes' rule

$$P(\text{cancer} \mid \text{smoker}) = P(\text{cancer}) . \frac{P(\text{smoker} \mid \text{cancer})}{P(\text{smoker})}$$

Posterior belief *Prior belief* *New knowledge*

Thomas Bayes

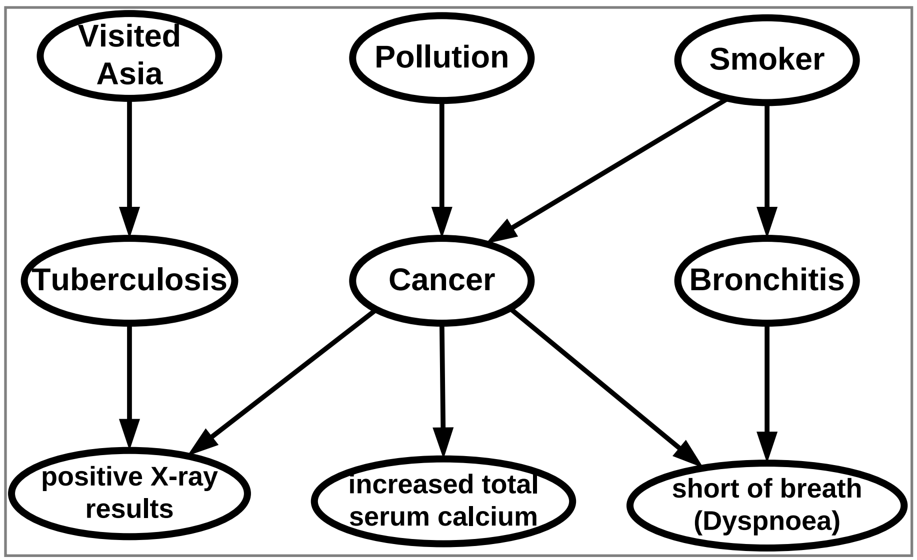
Types of questions

- **Probabilistic queries**
  - "I've got a temperature of 101, I'm a 37-year-old Male and my tongue feels kind of funny but I have no headache. What's the chance that I've got bubonic plague?"
- **Take decisions about tests**
  - Utility of tests: which tests give me maximal information
- **Take decisions** about interventions
- Explain things in terms of **causal relations**

➡ *How to represent this knowledge?*

Bayesian Networks

- Intuitive graphical representation expressing the relations among the variables. Can be causal relations.
- Plus probabilities attached to each node.



P(visited Asia)
P(pollution)
P(Smoker)
P(tuberculosis | visited Asia)
P(cancer |smoker , pollution)
P(Bronchitis|smoker)
...

Overview

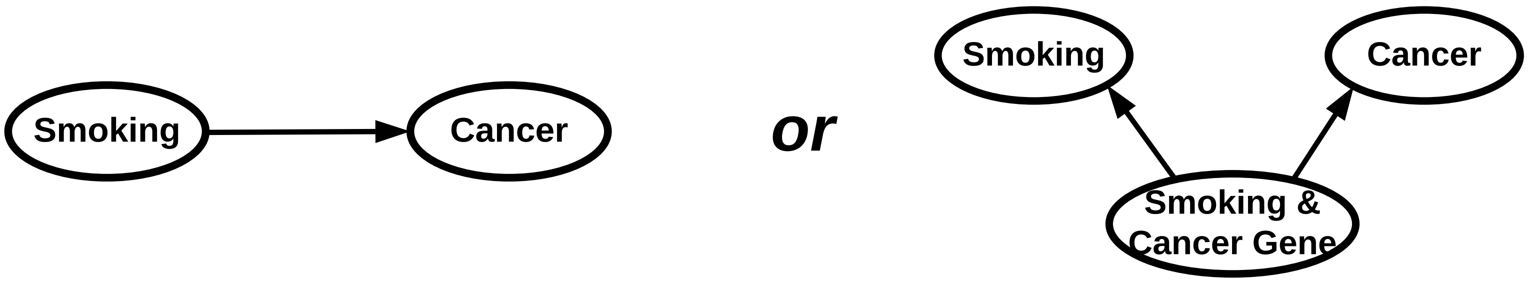
- Introduction: From Bayes' rule to Bayesian networks
- **The use of Bayesian networks**
- Learning from data
- Comparison with other techniques
- Conclusions

Typical use of Bayesian networks

- to model and explain a domain.
- to update beliefs about states of certain variables when some other variables were observed,
  - e.g.: P(car breaks down | age of car = 16, changed oil = no).
  - ≈ prediction
- to find most probable configurations of variables
- to support decision making under uncertainty (a Bayesian Network is a probabilistic model)
- to find good strategies for solving tasks in a domain with uncertainty

Typical use of Bayesian networks II

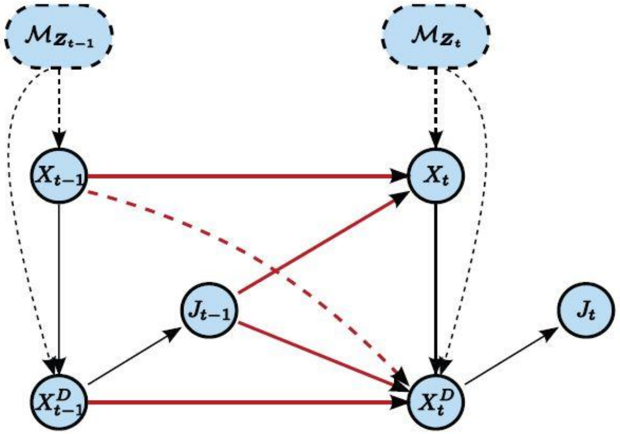
- Explain things in terms of causal relations
  - E.g.: smoking causes lung cancer?



- To answer qualitative questions
  - E.g.: if the national bank would lower interest rates, but the confidence remains low, would it help the economy?

Dynamic Bayesian Networks

- Models a dynamic system: the state at time ***t*** is affected by the state at time ***t-1***
- Used in reliability analysis



Graphical Duration Models

Rail defect

Examples

- Norsys (http://www.norsys.com/)
  - http://www.norsys.com/netlibrary/index.htm
  - *Coronary Risk*: a Bayesian Network to predict risk of Coronary Heart Disease
  - *Agricultural Yield*
  - *Car Diagnosis*
  - *Chest Clinic Decision*: A graphical method for solving a decision analysis problem
  - *Oil Wildcatter Extended*: decision network
  - *Win95pts*: An expert system for printer troubleshooting in Windows 95.

Up2Date technology

Overview

- Introduction: From Bayes' rule to Bayesian networks
- The use of Bayesian networks
- **Learning from data**
- Comparison with other techniques
- Conclusions

Vrije Universiteit Brussel

Learning Bayesian networks

From data and expert knowledge

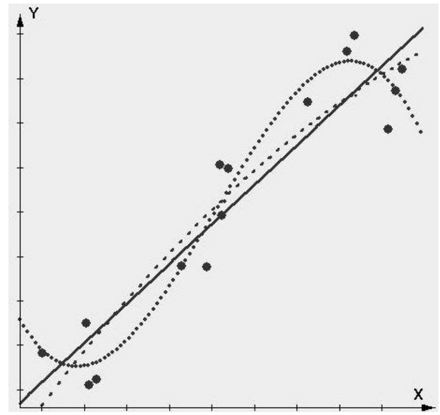
I. Parameterization ('the probabilities')
   - Based on assumptions on the relations
   - e.g. linear with Gaussian errors
   - When the structure is known

Learning Bayesian networks II

II. The structure of the graph

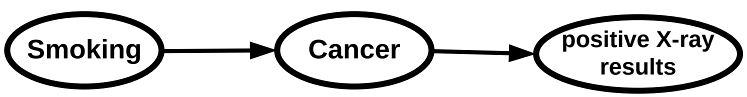
1. *Find minimal model that best explains the data*

Trade-off between goodness-of-fit and model complexity



overfitting in regression

2. *Find model that explains the conditional independencies*
   - A causal structure implies conditional independencies



➡ Causal structure is learned from data!!

Up2Date technology

Overview

- Introduction: From Bayes' rule to Bayesian networks
- The use of Bayesian networks
- Learning from data
- **Comparison with other techniques**
- Conclusions

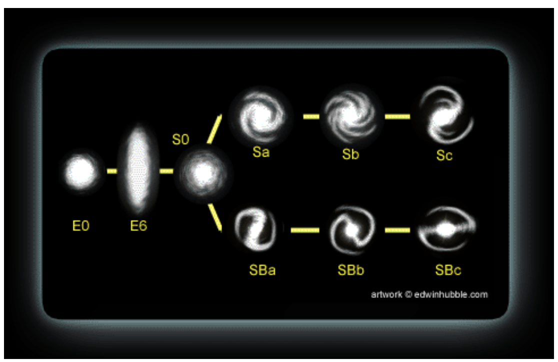
Vrije Universiteit Brussel

Statistical Learning

- **Goal**: learning from data, try to understand the underlying system that generated the data
- **Supervised learning**

A	B	C	D	E
2	12	0,42	TRUE	blue
1	73	1,93	FALSE	green
4	8	0,03	TRUE	red
2	27	2,84	TRUE	??

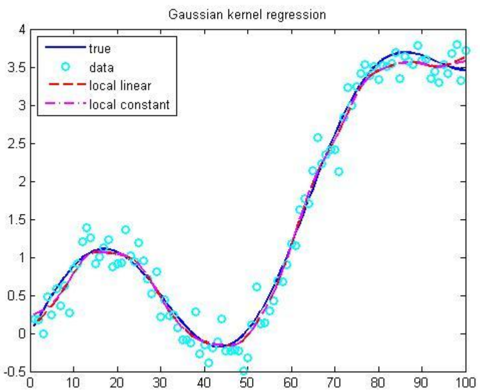
learn to predict E from A, B, C & D



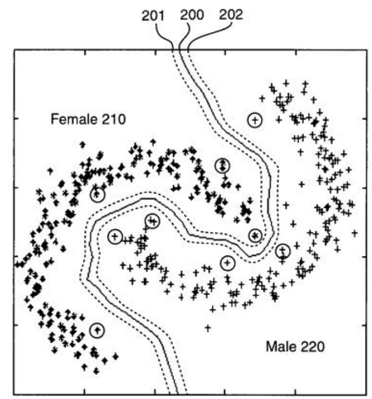
Classification of galaxies by Hubble telescope

Up2Date technology

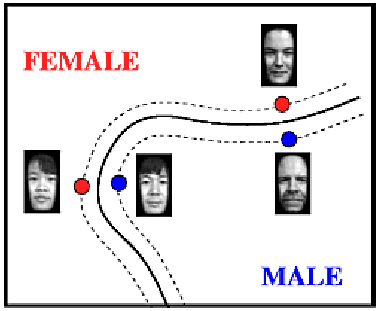
Techniques for supervised learning

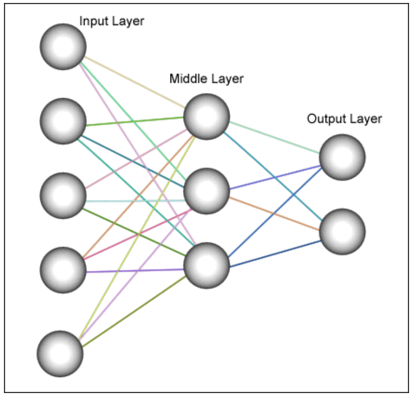


Regression



Support Vector Machines





Neural network

www.mastersininnovation.com

Vrije Universiteit Brussel

MASTERS IN INNOVATION

PICANOL why weavers win

P&G

RECTICEL

Unilever

ArcelorMittal

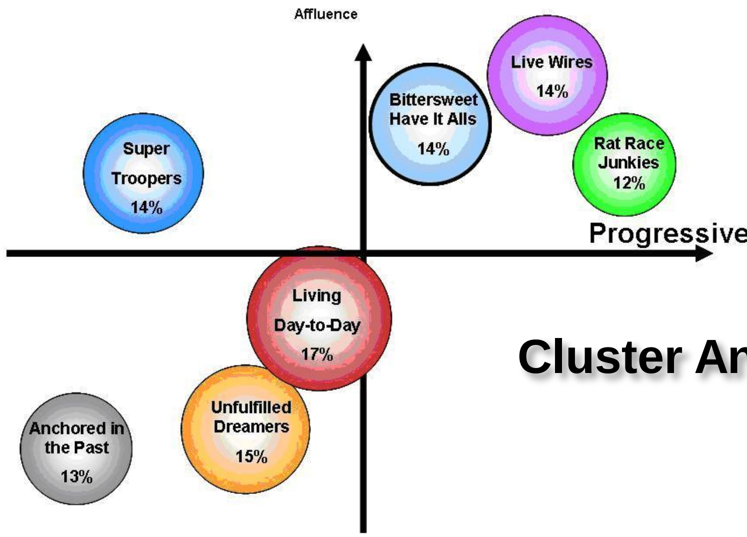
PHILIPS

VERHAERT

Statistical Learning II

- **Unsupervised learning**
  - there is no outcome measure
  - the goal is to describe the *associations* and *patterns* of the data

<u>Example</u>: find types of consumers



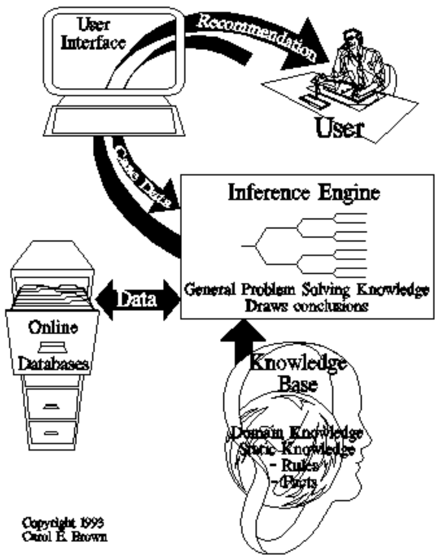
Cluster Analysis

Techniques for statistical learning compared

- Provide 'black box' models for the relation between the variables that we know and the ones we want to predict.
  - They are usually better in that task.
  - Bayesian networks model the relations among all variables.
  - Bayesian networks provide insight.
- Bayesian networks are not good in pattern recognition (e.g. recognizing characters from a printed text).

Up2Date technology

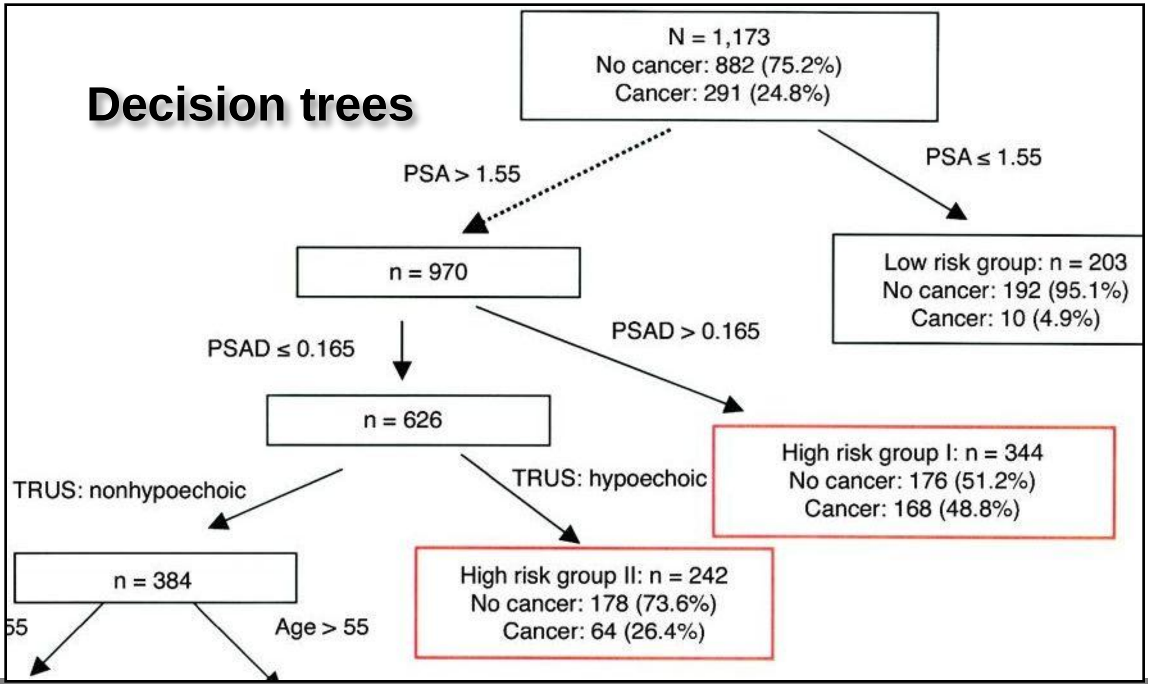
Knowledge representation



Expert Systems

Example: Evaluating car insurance risks.

Decision trees



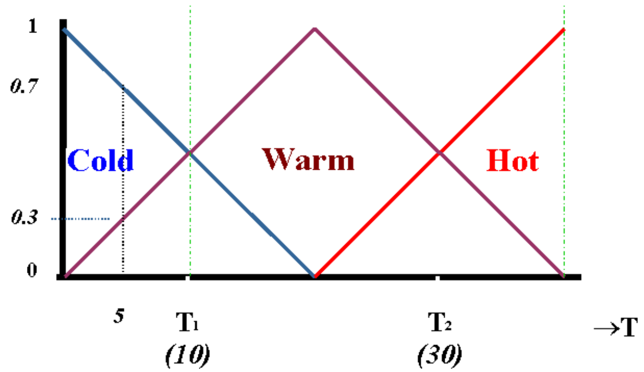
Rule-based systems

A => B
B & C => D

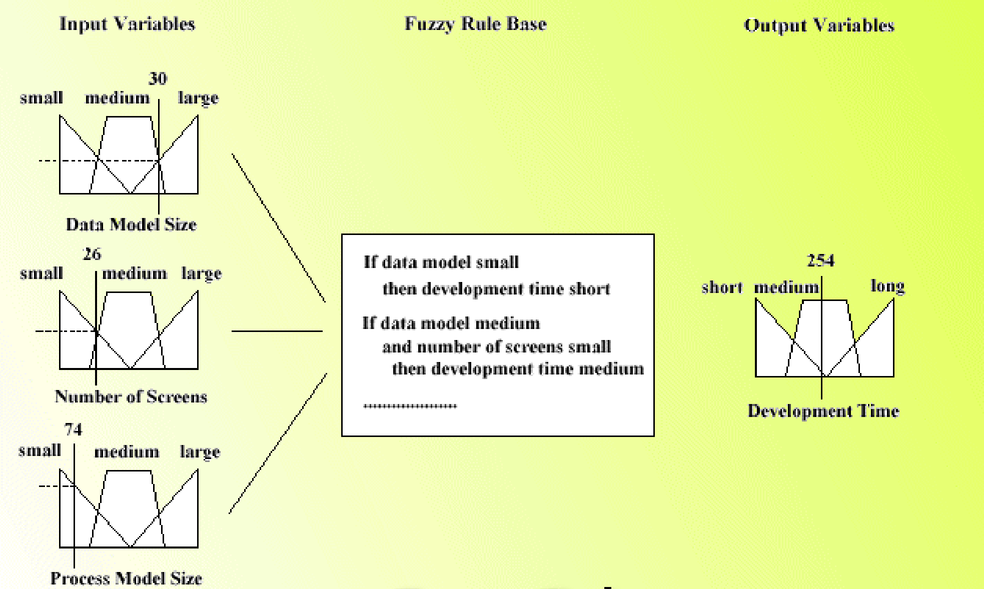
Up2Date technology

Knowledge representation II

www.mastersininnovation.com



Fuzzy Logic



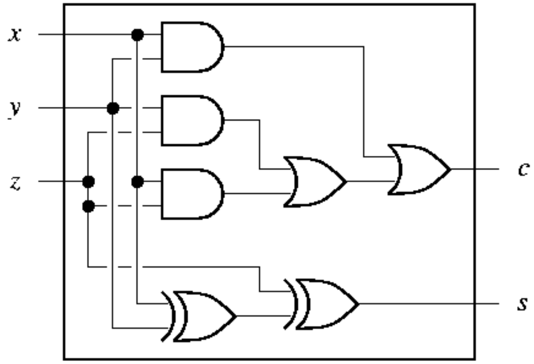
Fuzzy Rules

Vrije Universiteit Brussel

MASTERS IN INNOVATION — PICANOL why weavers win — P&G — RECTICEL — Unilever — ArcelorMittal — PHILIPS — VERHAERT

Knowledge representation techniques compared

- Bayesian networks can be regarded as the underlying (causal) structure, from which (fuzzy) rules can be extracted
- While the graph describes the relations among the variables, other techniques describe the *type* and *structure* of the relations better .



Digital Circuit

Overview

- Introduction: From Bayes' rule to Bayesian networks
- The use of Bayesian networks
- Learning from data
- Comparison with other techniques
- **Conclusions**

Conclusions

- A lot of data is available nowadays + a lot computing power
  - Data mining and statistical learning becomes interesting
  - Academic researchers are in need of real-world data and problems!
- A Bayesian network provides an intuitive graphical representation
  - Knowledge representation
- Learning algorithms exist that can learn the models from data
  - Extract knowledge from data

- **Useful when the relations among the variables matter**
- **Not when the (causal) relations are trivial**

References

- **My research**: http://parallel.vub.ac.be => research => causal inference
- **Norsys:** http://www.norsys.com/
  - http://www.norsys.com/netlibrary/index.htm (examples)
  - tutorials
- **Bayesia:** http://www.bayesia.com
  - Examples: http://www.bayesia.com/en/products/bayesialab/resources.php
- **Bayesian Network Repository**
  - http://compbio.cs.huji.ac.il/Repository/
- **Statistical Data Mining Tutorials**:
  - http://www.autonlab.org/tutorials/

Verhaert New Products & Services nv

Hogenakkerhoekstraat 21
9150 Kruibeke
Belgium
Tel +32 (0)3 250 19 00
Fax +32 (0)3 254 10 08
www.verhaert.com
info@verhaert.com