OVERVIEW

- Bayesian networks
  - advantages wrt the naïve Bayes model
  - construction of Bayesian networks

- Influence diagrams
  - advantages wrt decision trees
  - influence diagrams vs. (for) clinical practice guidelines
  - new model: decision analysis networks (DANs)

- Temporal PGMs
  - new models: event networks (non-Markovian) and dynamic LIMIDs (Markovian)
  - advantages wrt Markov decision trees
Elvira

- Research project of several Spanish universities
- Supported by national research agencies
- Elvira program
  - written in Java (advantage: portability; drawback: slowness)
  - ~120,000 lines of source code, publicly available on Internet:
    www.ia.uned.es/~elvira
  - advanced graphical interface for editing and evaluating models
    - in Spanish and English; easy to translate to other languages
  - several algorithms for inference and for learning from databases
  - weaknesses
    - still buggy
    - no on-line help yet
- Used for tuition and research in at least 8 countries

1. Bayesian networks
Old method: naïve-Bayes for probabilistic diagnosis

- $n$ diagnoses, $m$ variables representing possible findings
- 1st hypothesis: diagnoses are mutually exclusive (i.e., the patient has at most one disease)
- 2nd hypothesis: findings are conditionally independent

\[
P(f_1, \ldots, f_m | d_i) = P(f_1|d_i) \cdot \ldots \cdot P(f_m|d_i)
\]

- Bayes’ theorem (naïve method)

\[
P(d_i | f_1, \ldots, f_m) = \frac{P(f_1|d_i) \cdot \ldots \cdot P(f_m|d_i) \cdot P(d_i)}{\sum_j P(f_1|d_j) \cdot \ldots \cdot P(f_m|d_j) \cdot P(d_j)}
\]

\[
P(f_m | f_1, \ldots, f_{m-1}) = \frac{\sum_j P(f_1|d_j) \cdot \ldots \cdot P(f_{m-1}|d_j) \cdot P(d_j) \cdot P(f_m|d_j)}{\sum_j P(f_1|d_j) \cdot \ldots \cdot P(f_{m-1}|d_j) \cdot P(d_j)}
\]

Limitations of the naïve-Bayes method

- In general diagnoses are not mutually exclusive.
- In general findings are not conditionally independent.

- In the 70s, probability was discarded in artificial intelligence
- ... but came back in the 80s with Bayesian networks
Advantages of Bayesian networks (1/2)

- BNs are usually causal models
  - closer to doctors' reasoning: explanation of reasoning
  - probabilities are in general easier to obtain
- BNs can diagnose several diseases simultaneously
- BNs do not assume conditional independence
- BNs can be learnt from databases
- BNs can combine objective probabilities (frequencies) with subjective estimates
- Specific methods for sensitivity analysis in BNs
Advantages of Bayesian networks (2/2)

- Canonical models facilitate the construction of BNs
  - when the BN is built from human knowledge (subjective estimates)
  - and also when a BN is learnt from a database
    - Díez, Druzdzel. Canonical probabilistic models for knowledge engineering. 2005
- Canonical models lead to more efficient inference
  - Díez, Galán. Efficient computation for the noisy-MAX. 2003
- Several methods for the explanation of reasoning in BNs
  - useful for building and debugging Bayesian networks
    - Lacave, Onisko, Díez. Use of Elvira’s explanation facility for debugging. 2006.
  - useful for avoiding human reluctance to accept expert systems
  - useful for using BNs as tutoring systems (e.g. for students of medicine)

Use of BNs in real world applications

- BNs are more and more popular in artificial intelligence, not only in Academy but also in industry
- Many applications:
  - medicine: diagnostic expert systems
  - genetics: modeling gene interactions
  - epidemiology: detecting and quantifying causal influences
    - Program on Causal Inference in Epidemiology (Harvard; director: J. Robins)
  - agriculture, computer security, e-commerce, etc., etc.
- In contrast, BNs are almost unknown in medicine
  - Textbooks only describe the naïve Bayes method (and quite superficially, by the way)
- Why?
2. Influence diagrams

A medical problem

- **Disease** $X$
  - Prevalence: $P(+x) = 0.14$

- **Therapy** $D$
  - Utility:
    
    | $u(x, d)$ | $+x$ | $-x$ |
    |-----------|------|------|
    | $+d$      |  8   |  9   |
    | $-d$      |  3   | 10   |

- **Test** $Y$
  - Sensitivity: $P(+y|+x) = 0.91$
  - Specificity: $P(-y|-x) = 0.97$
  - Cost: $u_{test}(x, d) = u_{not-test}(x, d) - 0.2$

- **Decisions:**
  - Is it worthy to do the test?
  - In what cases should we apply the therapy?
F. J. Diez and M. Luque

Probabilistic graphical models for medical decision making

ESF-IfW Conference on The Global Health Economy, 2006
Advantages of influence diagrams (1/3)

- IDs are more compact than decision trees
  - An ID having $n$ binary nodes ~ a DT having $2^n$ branches

- Explicit representation of causality
  - A link indicates causal influence
  - The absence of a link means “no causal influence” (hypothesis)

- IDs are much easier to build than decision trees
  - IDs use direct probabilities (prevalence, sensitivity, specificity...) and costs (mortality, morbidity, economic cost...)
  - No external pre-calculation of probabilities is required
  - IDs can use super-value nodes: explicit combination of utilities
  - Each parameter appears only once in the ID
    - In many cases it is not necessary to have parametric variables
Advantages of influence diagrams (2/3)

- Having all the information, no debugging is usually needed
  - On the contrary, “all trees have bugs” (Primer on MDA)
- Parametric sensitivity analysis is much easier
- IDs are much easier to modify than decision trees
  - Refine the model with new decisions and chance variables
  - Structural sensitivity analysis is incomparably easier
  - Can adapt to different regional settings
  - Can adapt to patient's medical characteristics and preferences
- IDs transform automatically into decision trees
  - ... but the reverse is not true (no general algorithm)
  - If you build a decision tree, you only have a decision tree.
  - If you build an ID, you have both, with much less effort.

Advantages of influence diagrams (3/3)

- Two possibilities of evaluation:
  1. expansion of an equivalent decision tree
     - exponential complexity (time and space)
     - equivalent to the brute-force method for Bayesian networks
     - many problems can not be solved by this method
  2. operations on the ID (recursive reduction of the ID)
     - direct manipulation of the graph and/or potentials of the ID
     - similar to the best algorithms for Bayesian networks
     - canonical models and SV nodes can lead to more efficient evaluations
- More possibilities of explanation of reasoning
  - computation of posterior probabilities on the ID (as if it were a BN)
  - value of information (EVPI and other measures) can be computed easily
  - other methods from Bayesian networks and qualitative prob. networks.
  - These methods can be used for debugging/refining IDs.
Clinical practice guidelines (CPGs)

- Construction of CPGs
  - Usually: expert opinion or consensus of experts
  - Another possibility: influence diagrams
    - Sanders, Nease, Owens: several papers on building CPGs from IDs.

- Advantages of an ID wrt a CPG
  - explicit decision model
    - easily combine expert opinions and evidence (statistical data)
    - help in difficult cases, in which the policy is not evident
  - flexibility: can be extended and adapted, as mentioned above
  - can include patients’ preferences
  - the physician plays an active role,
    he/she is not a passive user of CPGs developed by others

A proverb

- Don’t give a man a fish; give him a rod and teach him how to fish.
- Don’t give a doctor a written CPG; give them an influence diagram and teach them how to use Elvira.
IDs in the literature on MDM

- Journal: *Medical Decision Making*
  - very few papers using IDs in their analyses

- Books that mention decision trees and do not mention IDs
    [Influence diagrams were first published in (Howard and Matheson, 1984)]
  - Sacket et al. *Evidence-Based Medicine*. 1997
    (and three other books on EBM).

IDs in the literature on MDM (cont.)

Books that mention decision trees and do not mention IDs (cont.)

- One book that mentions IDs
  - Chapman, Sonnenberg (eds.). *Decision Making in Health Care*. 2000
    (5 pages out of 421).

- Another book that mentions IDs
    “An influence diagram (also known as a tornado diagram) ...” [p. 242]

- Conclusion: informal survey of books on MDM and EBM
  - 10 books on MDM and several on EBM published after 1984
  - All of them mention DTs but only one mentions IDs, quite briefly
Limitations of IDs

- Dealing with asymmetric problems
  - Standard IDs are symmetric
  - Some software tools (e.g., TreeAge) allow asymmetry
    - but sometimes “arcs of asymmetry” are not intuitive
  - Many asymmetric problems can not be solved with IDs

- Limitations of current software packages
  - Very few packages allow sensitivity analysis directly on IDs.
  - No package allows cost-effectiveness analysis directly on IDs.

- Solutions
  - More powerful software tools (e.g., future versions of Elvira)
  - More flexible representation models for asymmetric problems

Decision-analysis networks (DANs)

- Very similar to IDs, but:
  - DANs do not have information arcs
  - DANs do not require a total ordering of decisions
  - Some nodes are marked as “always known” (for instance, symptoms)
  - DANs may have revelation arcs: “Dec:Test”→“Result of test”
3. Temporal PGMs

**Temporal PGMs**

- **Non-Markovian models**
  - For instance, birth delivery happens 9 months after conception
  - New model: networks of events (temporal Bayesian networks)

- **Markovian models**
  - Influence diagrams with Markov nodes
    - A node in an ID that represents a (small) Markov model
  - Other models: factored MDPs, factored POMDPs…
  - New model: 2TLIMIDs
2TLIMID for a simplified medical example

- It would be difficult to build a Markov decision tree for this problem.

2TLIMID for a real-world example

- It would be impossible to build a Markov decision tree for this problem.
Conclusion

◆ Advantages of PGMs
  ➢ Bayesian networks vs. naïve Bayes method
  ➢ Influence diagrams vs. decision trees
  ➢ Influence diagrams vs. (for) clinical practice guidelines
  ➢ Temporal PGMs (2TLMIDs, etc.) vs. Markov decision trees

◆ Nevertheless, PGMs are almost unknown in medicine

◆ Our research
  ➢ new types of models (representation)
  ➢ algorithms for “old” and new models
  ➢ public software tool, Elvira (www.ia.uned.es/~elvira)
  ➢ real-world models for several medical problems

◆ Our interest
  ➢ collaborating with other groups doing research on medical decision analysis