Harvard School of Public Health
Boston, 16 October 2002

Elvira

A software package for probabilistic DAGs

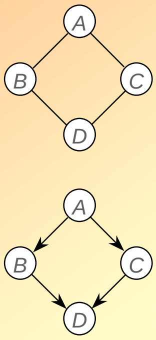
www.ia.uned.es/~elvira

F. Javier Díez
Dept. Artificial Intelligence. UNED
Madrid, Spain
www.ia.uned.es/~fjdiez

OVERVIEW

- Introduction:
  - graphs
  - Bayesian networks: definition, semantics, construction
- Elvira
  - the project
  - probabilistic reasoning with Elvira:
    conditional independence, *d*-separation, the Markov property
  - real-world examples: Prostanet, Nasonet, Hepar II
- Causal Bayesian networks for epidemiological research
  - use of Elvira as a pedagogical tool
- Conclusions

Graphs

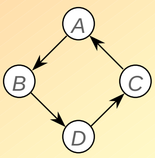


- Graph: nodes + links
- Undirected link: (A, B) is the same as (B, A)
  Directed link: (A, B) is not the same as (B, A)
- Undirected graph: all the links are undirected
  Directed graph: all the links are directed
- Cycle (in an undirected graph): any closed path
- Cycle (in a directed graph): a closed path that follows the directions of the links

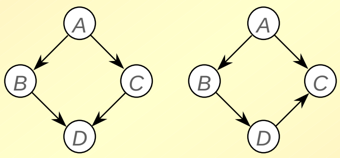
Graphs → Undirected graphs
Graphs → Directed graphs
Directed graphs → Acyclic directed graphs
Directed graphs → Cyclic directed graphs

Acyclic directed graphs

- Cycle

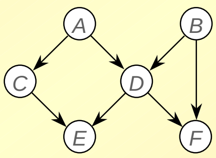


- Loops

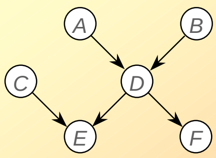


- Acyclic directed graphs

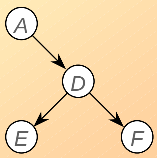
multiply-connected



polytree



tree



Bayesian network

- Elements
  - Set of variables $\{X_i\}$
  - Acyclic directed graph
    - Each node in the graph represents a variable X_i
  - Conditional probability distribution for each variable: $P(x_i \mid pa(x_i))$
    - For a node without parents: $P(x_i \mid pa(x_i)) = P(x_i)$
- Result: join probability distribution

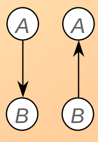
$$P(x_1, \cdots, x_n) = \prod_{i=1}^{n} P(x_i | pa(x_i))$$

- Markov property
  - Given a node X_i and a set of variables $\{Y_j\}$ such that no Y_j is a descendant of X_i in the graph, it hods that

$$P(x_i \mid pa(x_i), y_1, \ldots, y_n) = P(x_i \mid pa(x_i))$$

Semantics of a Bayesian network

- As a mathematical model (by definition): graph + joint prob
  - It represents probabilistic dependencies and independencies
  - A link, taken individually, has no meaning
  - Two networks are (Markov) equivalent if they represent the same set of dependencies and independencies
- As a representation of the real world
  - Causal links are those that represent real-world mechanisms and influences
  - Causal and non-causal links may coexist in the same network
  - A network is causal when all of its links are causal
  - A Bayesian network may be causal…
  - … but it is not strictly necessary
  - Two different causal networks are never equivalent, even though they represent the same (in)dependencies.



Three ways to build a Bayesian network

- From a database

  Data base $\xrightarrow{\text{algorithm}}$ Bayesian network

  - There are many algorithms, every year many new algorithms
  - Similar to statistical methods (logistic regression, neural nets...)
- With causal knowledge from human experts

  Known causal mechanisms $\xrightarrow{\text{modeling}}$ Causal graph $\xrightarrow{\text{probabilities}}$ Bayesian network

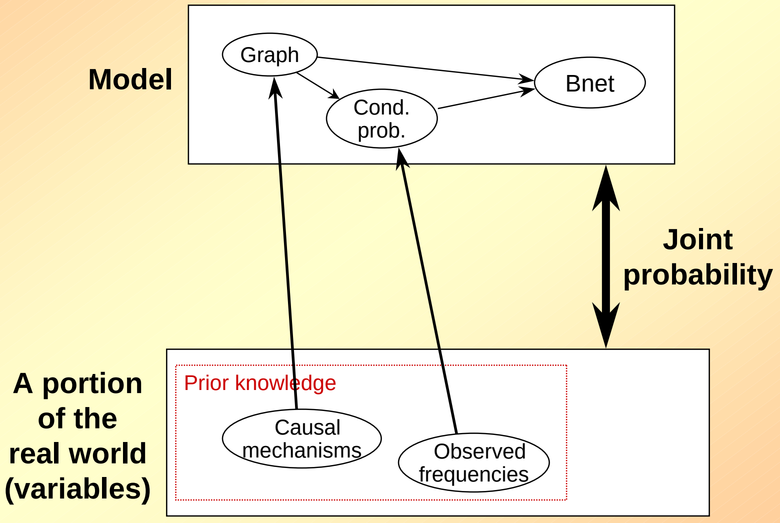
  - (roughly) Each link represents a causal mechanism
  - It would be impossible to build a BN without causal knowledge
- Hybrid method:
  - experts $\rightarrow$ graph; database $\rightarrow$ probabilities

A modified version of the causal Markov principle

- If a graph is
  - causal
  - complete (in a causal sense)
  - detailed enough
  - acyclic

  then the graph satisfies the (probabilistic) Markov condition with respect to the real-world probability distribution
- Causal: each link represents a causal mechanism
- Complete: for each causal mechanism, there is a path not containing extraneous variables
- Detailed enough:
  - granularity of each variable
  - intermediate causes
- These conditions guarantee the Markov property

The principle and its application

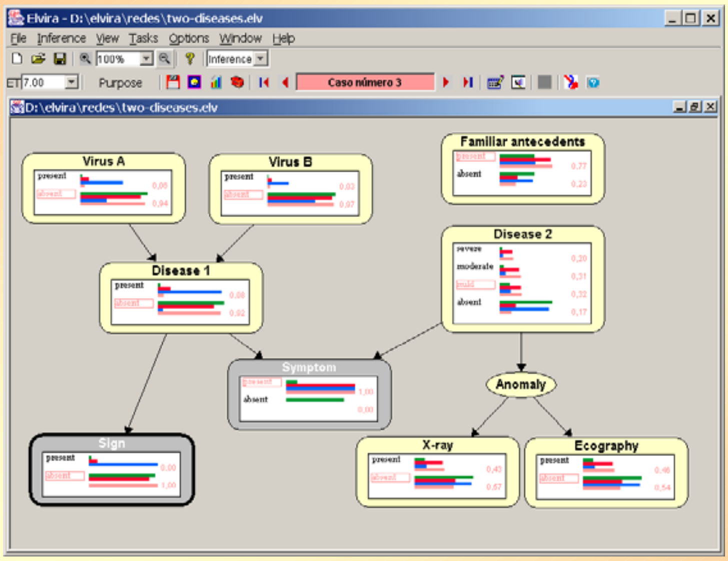


Elvira

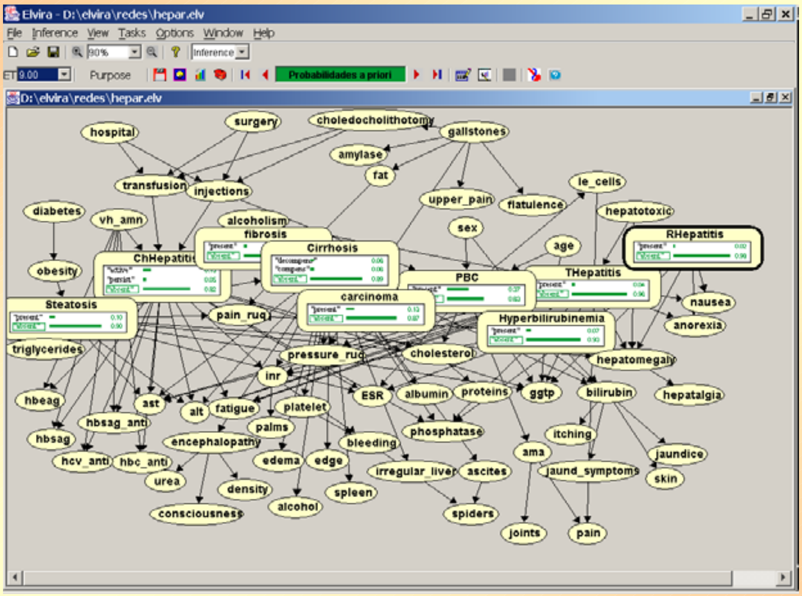
- Research project of several Spanish universities
- Supported by the government research agencies
  - Elvira I (1997-2000)
  - Elvira II (2001-2004)
- Elvira program
  - written in Java (advantage: portability; drawback: slowness)
  - 115.000 lines of source code, publicly available on Internet: `www.ia.uned.es/~elvira`
  - 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
    - poor documentation of source code

Medical diagnosis with Elvira

- Bayes theorem
- Independence, correlation and causal graphs
  - correlated variables
  - conditional independence
  - conditional dependence
  - *d*-separation in Bayesian networks
- Probabilistic diagnosis
  - weighting different findings
    - convergent findings
    - contradictory findings
  - differential diagnosis
    - explaining away
    - diagnosis by exclusion

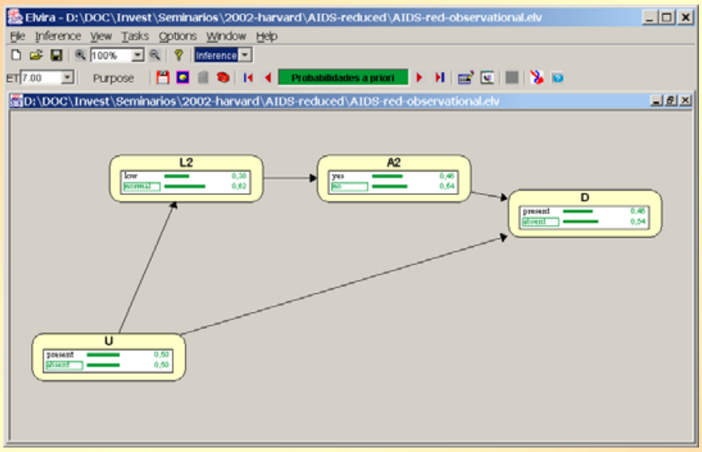


A simplified medical example that illustrates *d*-separation



HEPAR, a real-world Bayesian network for liver diseases

Bayesian networks
for epidemiological research



Conditioning on A_2 opens a non-causal backdoor path from A_2 to D that confounds the causal effect $A_2 \rightarrow D$

Conclusion: utility of Elvira

- Tool for building Bayesian networks and influence diagrams
  - Examples: Prostanet, Nasonet, Hepar II, etc.
- Pedagogical tool
  - illustrate *d*-separation and the Markov property
  - illustrate the problems of analyzing causality given observational data
- Future: tool for epidemiological studies
  - input: causal graph + observational database
  - first step: eliminate unmeasured variables from the graph result: statistical graph
  - second step: obtain the conditional probabilities for the statistical graph from the database
  - output: causal risk ratios