Massachusetts Institute of Technology
Boston, October 13, 2015

Probabilistic graphical models in artificial intelligence and medicine

Francisco Javier Díez
Dept. Artificial Intelligence. UNED
Madrid, Spain

www.ia.uned.es/~fjdiez
www.cisiad.uned.es

OVERVIEW

- Probability in artificial intelligence
- The naïve-Bayes method
- Influence diagrams
- Decision analysis networks
- Cost-effectiveness analysis
- Markov models
- Conclusion

Probability in artificial intelligence

- A.I. was "born" in 1956, at the Dartmouth Conference
- In the first 25 or 30 years, many researchers doubted or denied that probability could play a significant role in A.I.
- First reason (cf. [Sutton and Barto, 1998]):
  - Computers were already good at arithmetic operations
  - but could not perform "easy" tasks (easy for a little child): vision (image understanding), natural language, planning…
  - Those tasks could not be solved with arithmetic operations; they require conceptual reasoning (symbol manipulation → LISP).
  - Probabilistic "reasoning" consisted mainly in number crunching, not in conceptual reasoning.
- Second reason: limitations of probabilistic methods.

Naïve-Bayes method for probabilistic diagnosis

- *n* diagnoses, *m* possible findings
- 1st hypothesis: diagnoses are <u>mutually exclusive</u>
  (i.e., the patient has at most one disease)
- 2nd hypothesis: findings are <u>conditionally independent</u>

$$P(f_1, \ldots, f_m \mid d_i) = P(f_1 | d_i) \cdot \ldots \cdot P(f_m | d_i)$$

- Bayes' theorem (naïve method)

$$P(d_i | f_1, \ldots, f_m) = \alpha \cdot P(f_1 | d_i) \cdot \ldots \cdot P(f_m | d_i) \cdot P(d_i)$$

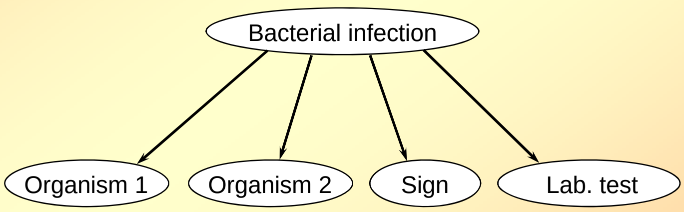
Succesfull applications of the naïve-Bayes

- Lodwick GS, Haun CL, Smith WE, et al. , “Computer diagnosis of primary bone tumors: A preliminary report,” *Radiology* **80** (1963) 273-275.
- Overall JE, Williams CM, “Conditional probability program for diagnosis or thyroid function,” *JAMA* **183** (1963) 307-313.
- Toronto AF, Veasy LG, Warner HR, “Evaluation of a computer program for diagnosis of congenital heart disease,” *Progress in Cardiovascular Diseases* **5** (1963) 362-377.
- Warner HR, Toronto AF, Veasy LG, “Experience with Bayes’ theorem for computer diagnosis of congenital heart disease,” *Annals New York Acad. Sciences* **115** (1964) 558-567.
- de Dombal FT, Leaper JR Staniland JR, et al., “Computer-aided diagnosis of acute abdominal pain,” *BMJ* **2** (1972) 9—13.
- Gorry GA, Kassirer JP, Essig A, Schwartz WB, “Decision analysis as the basis for computer-aided management of acute renal failure,” *Amer. J Med* **55** (1973) 473-484.
- Gorry GA, Silverman H, Pauker SG, “Capturing clinical expertise: A computer program that considers clinical responses to digitalis,” *Amer. J. Med* **64** (1978) 452-460.

Some approximations were necessary for the sequential selection of tests [Gorry and Barnet, 1968].

Limitations of the naïve-Bayes method

- In general the diagnoses are not mutually exclusive: how to diagnose multiple disorders.
- In general findings are not conditionally independent.



ARTIFICIAL INTELLIGENCE 11 (1978) 115-144 115

Categorical and Probabilistic Reasoning in Medical Diagnosis*

Peter Szolovits, Ph.D.

Clinical Decision Making Group, Laboratory for Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.

Stephen G. Pauker, M.D. **

Department of Medicine, New England Medical Center Hospital, Tufts University School of Medicine, Boston, MA 02111, U.S.A.

Recommended by N. S. Sridharan

ABSTRACT

Medical decision making can be viewed along a spectrum, with categorical (or deterministic) reasoning at one extreme and probabilistic (or evidential) reasoning at the other. In this paper we examine the flowchart as the prototype of categorical reasoning and decision analysis as the prototype of probabilistic reasoning. Within this context we compare PIP, INTERNIST, CASNET, and MYCIN—four of the present programs which apply the techniques of artificial intelligence to medicine. Although these systems can exhibit impressive expert-like behavior, we believe that none of them is yet capable of truly expert reasoning. We suggest that a program which can demonstrate expertise in the area of medical consultation will have to use a judicious combination of categorical and probabilistic reasoning—the former to establish a sufficiently narrow context and the latter to make comparisons among hypotheses and eventually to recommend therapy.

Limitations of probability for AI in medicine

P. Szolovits. *Artificial Intelligence in Medicine*. Westview Press, 1982.

"The chief disadvantages of the decision theoretic approach are the difficulties of obtaining reasonable estimates of probabilities and utilities for a particular analysis. Although techniques such as sensitivity analysis help greatly to indicate which potential inaccuracies are unimportant, the lack of adequate data often forces artificial simplifications of the problem and lowers confidence in the outcome of the analysis. Attempts to extend these techniques to large medical domains in which multiple disorders may co-occur, temporal progressions of findings may offer important diagnostic clues, or partial effects of therapy can be used to guide further diagnostic reasoning, have not been successful. The typical language of probability and utility theory is not rich enough to discuss such issues, and its extension within the original spirit leads to untenably large decision problems. […]

A second difficulty for decision analysis is the relatively mysterious reasoning of a decision theoretic program—an explanation of the results is to he understood in terms of the numeric manipulations involved in expected value computations, which is not a natural way of thinking for most people."

Bayesian networks

Early publications on BNs

- Pearl J. "Reverend Bayes on inference engines: A distributed hierarchical approach", *Proc. AAAI*, 1982, Pittsburgh, PA, pp. 133-136.
- Kim J, Pearl J. "A computational model for combined causal and diagnostic reasoning in inference systems", *Proc. IJCAI*, pp. 190-193, 1983.
- Cooper G. *NESTOR: A Computer Based Medical Diagnostic Aid that Integrates Causal and Probabilistic Knowledge*, Ph.D. dissertation, Stanford Univ., 1984.
- Pearl J. "How to do with probabilities with people say you can't", *2nd Conference on AI Applications*, Miami, FL, 1985.
- Pearl J. "Fusion, propagation and structuring in belief networks". *AI* **29** (1986) 241-288.
- Pearl J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, San Mateo, CA: Morgan Kaufmann, 1988.
- Lauritzen SL, Spiegelhalter DJ. "Local computations with probabilities on graphical structures and their application to expert systems". *J. Royal Stat. Soc. B* **50** (1988) 157-224.
- Andreassen S, Woldby M, Falck B et al. MUNIN–A causal probabilistic network for interpretation of electromyographic findings. *Proc. IJCAI*, pp. 366-372, 1987.
- Heckerman D.E. *Probabilistic Similarity Networks*. Ph.D. dissertation, Stanford Univ., 1990. Published as a book: MIT Press, 1991.

BNs vs. naïve Bayes

- BNs can diagnose several diseases simultaneously
- BNs do not assume conditional independence
- Three types of reasoning:
  - abductive
  - deductive
  - inter-causal

OpenMarkov. Main features

- Strengths
  - Written in Java: portability (Windows, linux, MacOS…)
  - Open source
  - Software engineering tools: JUnit, maven, mercurial (bitbucket), nexus, bugtracker, etc.
  - Easily extensible: users can adapt it to their needs
  - Many types of models, potentials, etc.
  - Very active: new features are continuously added
  - Support for users and developers: wiki, lists, mail…
  - Well-documented format for encoding networks: ProbModelXML.
- Weaknesses
  - Written in Java: relatively slow (in some cases)
  - No on-line help, documentation still poor
  - Still a prototype; needs debugging
  - Support is limited, due to scarcity of human resources.

OpenMarkov
Español | English

Home
Users / Download
Developers
References
Acknowledgments
News

OpenMarkov

OpenMarkov is a software tool for probabilistic graphical models (PGMs) developed by the Research Centre for Intelligent Decision-Support Systems of the UNED in Madrid, Spain.

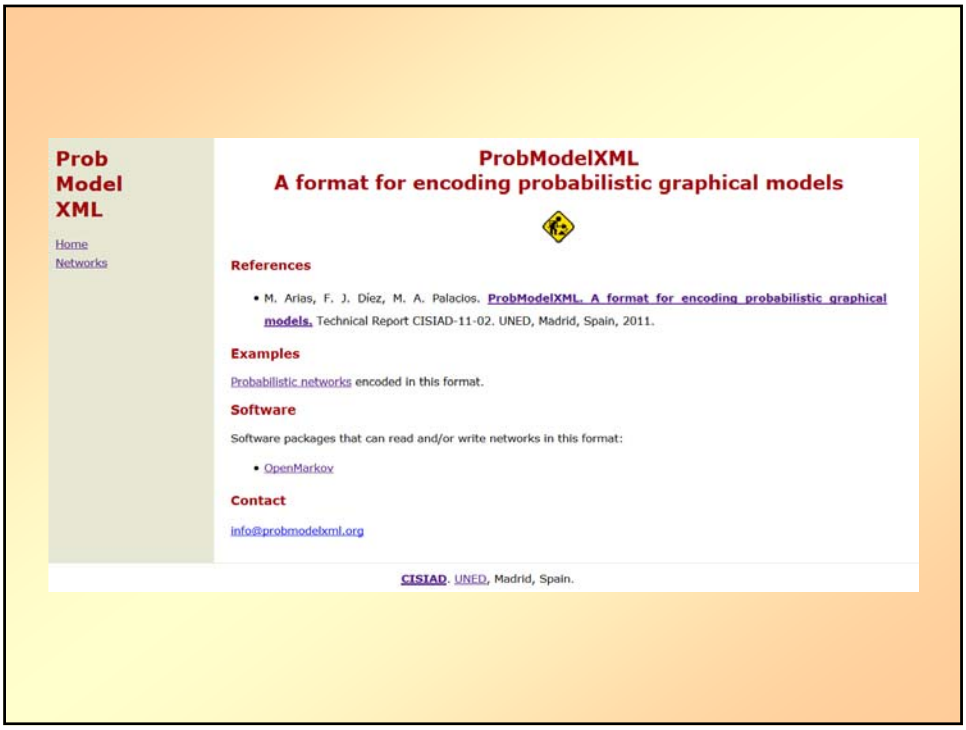
It has been designed for:

- editing and evaluating several types of several types of PGMs, such as Bayesian networks, influence diagrams, factored Markov models, etc.;
- learning Bayesian networks from data interactively;
- cost-effectiveness analysis.

You can read the tutorial to have a glimpse of its capabilities.

Visit the **users' page** to download **OpenMarkov** and obtain additional information.

CISIAD. Research Center on Intelligent Decision-Support Systems. UNED. Madrid, Spain.



Prob
Model
XML

Home
Networks

ProbModelXML
A format for encoding probabilistic graphical models

References

- M. Arias, F. J. Díez, M. A. Palacios. **ProbModelXML. A format for encoding probabilistic graphical models.** Technical Report CISIAD-11-02. UNED, Madrid, Spain, 2011.

Examples

Probabilistic networks encoded in this format.

Software

Software packages that can read and/or write networks in this format:

- OpenMarkov

Contact

info@probmodelxml.org

CISIAD. UNED, Madrid, Spain.

The MADP Toolbox 0.3.1

Frans A. Oliehoek
University of Amsterdam, Amsterdam, The Netherlands
University of Liverpool, Liverpool, United Kingdom

Matthijs T. J. Spaan
Delft University of Technology
Delft, The Netherlands

Philipp Robbel
Massachusetts Institute of Technology
Cambridge, MA, USA

João V. Messias
University of Amsterdam, Amsterdam, The Netherlands

April 10, 2015

Abstract

This is the user and developer guide accompanying the version 0.3.1 release of the Multiagent Decision Process (MADP) Toolbox. It is meant as a first introduction to the organization of the toolbox, and tries to clarify the approach taken to certain implementation details. In addition, it covers a few typical use cases and provides an installation guide. This document complements the automatically generated API reference.

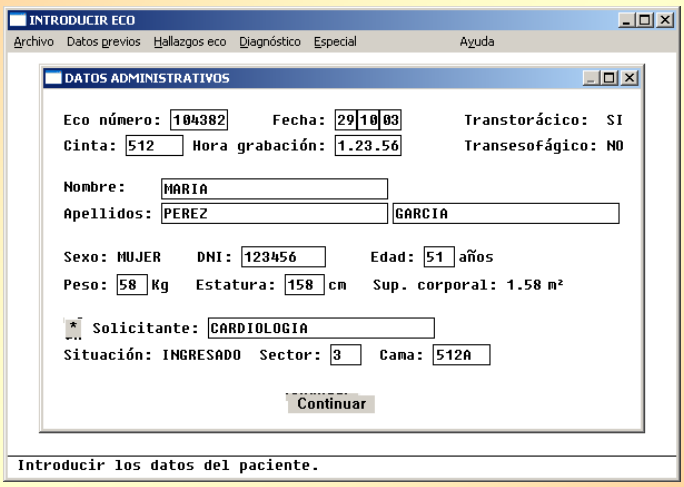
Contents

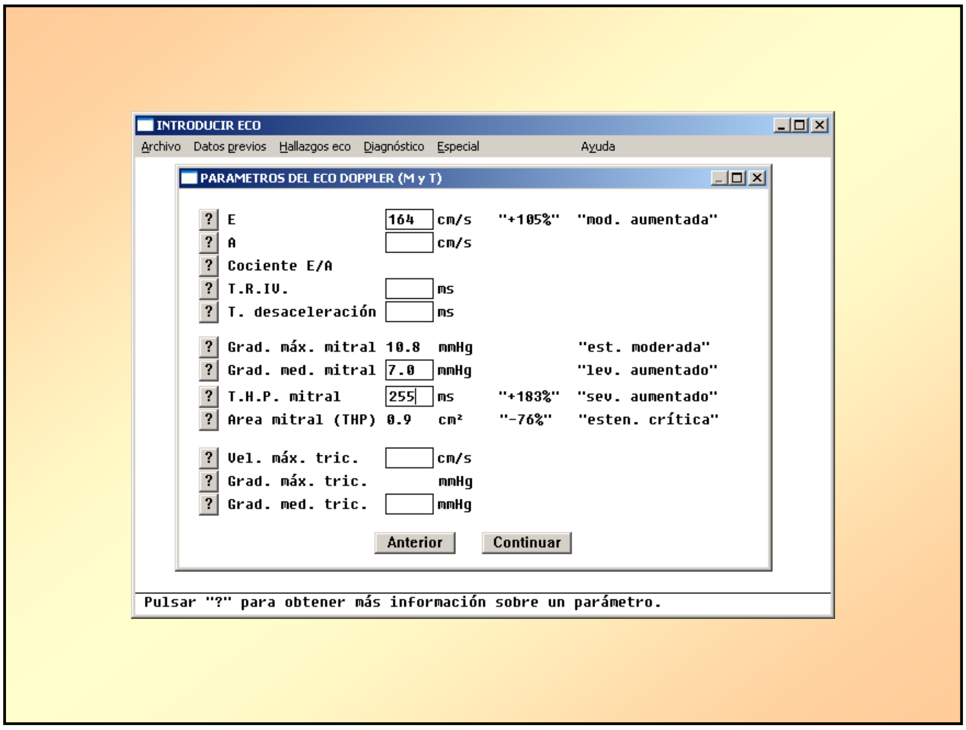
1

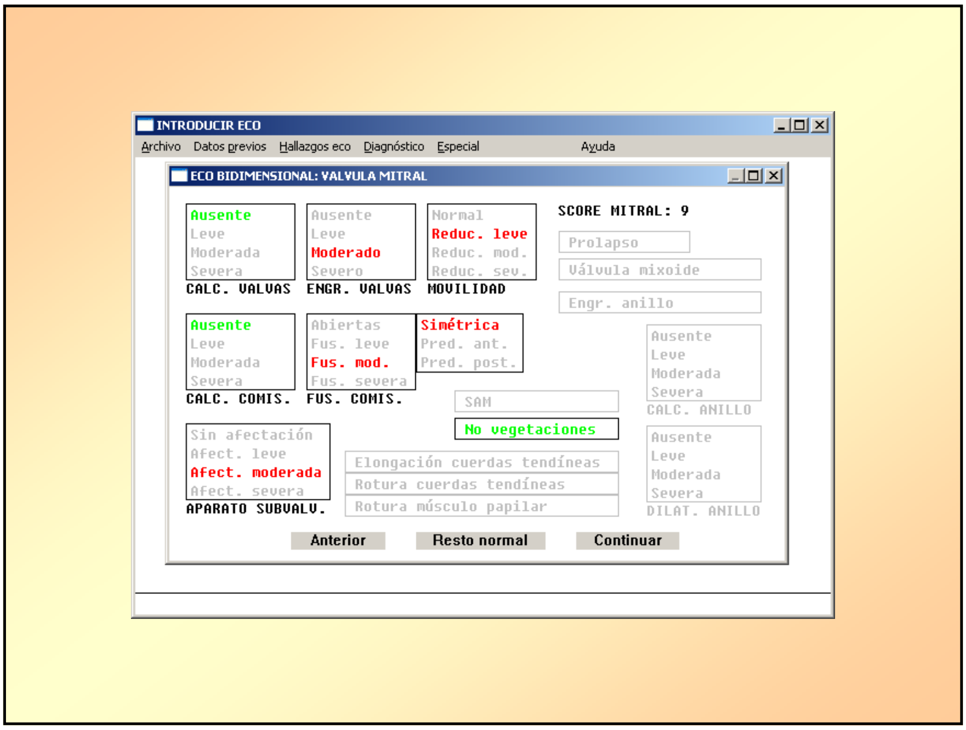
Our research on Bayesian networks (1/4)

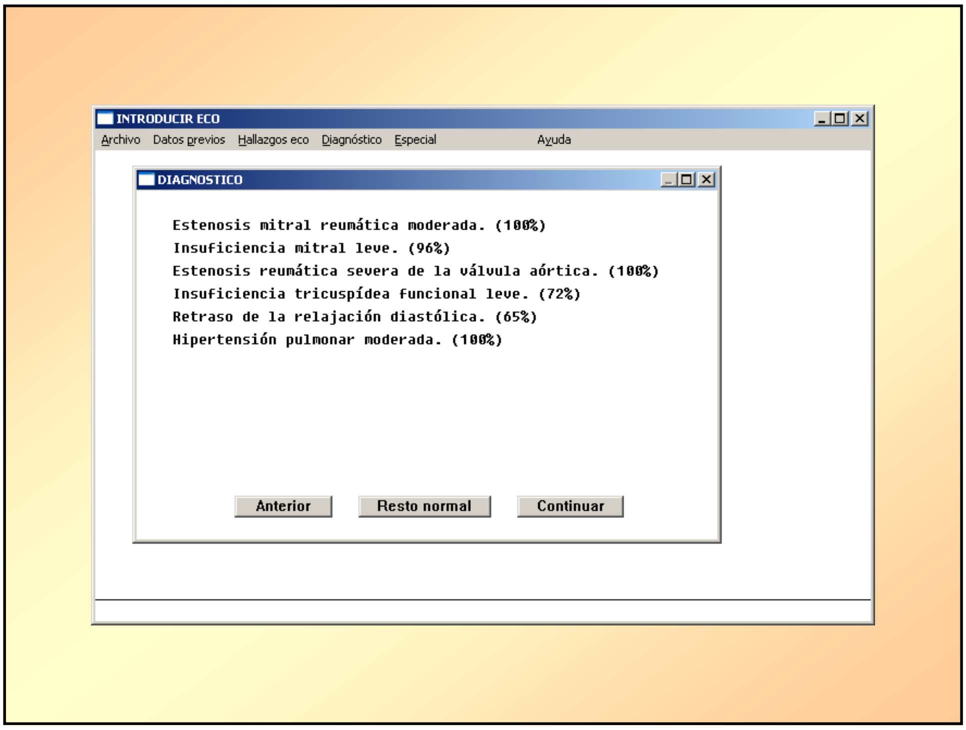
- Medical Bayesian networks we have built
  - DIAVAL: echocardiography (valvulopathies)
    F. J. Díez' thesis, 1994
  - Prostanet: urology (prostate cancer)
    Carmen Lacave's thesis, 2003
  - Nasonet: nasopharyngeal cancer spread
    Severino Galán's thesis, 2003
  - HEPAR II: liver dieseases
    Agnieszka Onisko's thesis, 2003
  - Catarnet: Cataract surgery
    Nuria Alonso's thesis, 2009

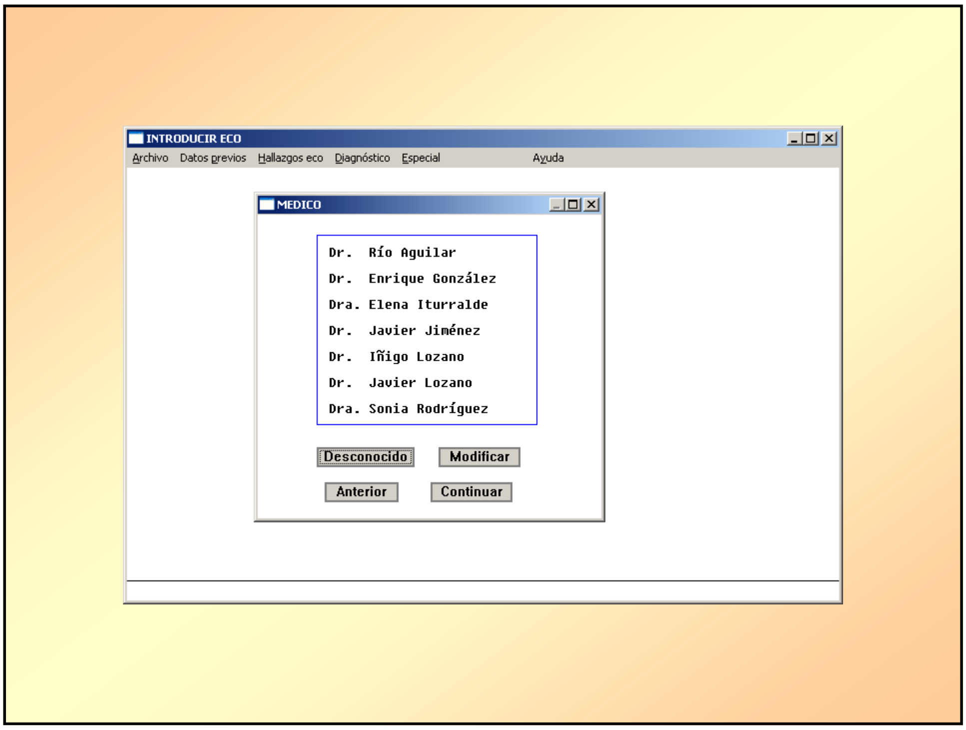
DIAVAL

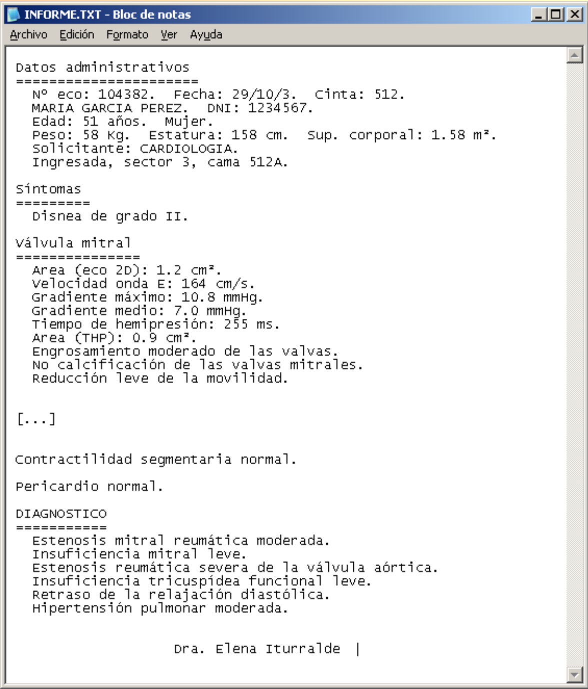



INTRODUCIR ECO
Archivo Datos previos Hallazgos eco Diagnóstico Especial Ayuda
PARAMETROS DEL ECO DOPPLER (M y T)
E 164 cm/s "+105%" "mod. aumentada"
A cm/s
Cociente E/A
T.R.IV. ms
T. desaceleración ms
Grad. máx. mitral 10.8 mmHg "est. moderada"
Grad. med. mitral 7.0 mmHg "lev. aumentado"
T.H.P. mitral 255 ms "+183%" "sev. aumentado"
Area mitral (THP) 0.9 cm² "-76%" "esten. crítica"
Vel. máx. tric. cm/s
Grad. máx. tric. mmHg
Grad. med. tric. mmHg
Anterior
Continuar
Pulsar "?" para obtener más información sobre un parámetro.


INTRODUCIR ECO
Archivo Datos previos Hallazgos eco Diagnóstico Especial Ayuda
ECO BIDIMENSIONAL: VALVULA MITRAL
Ausente Leve Moderada Severa
CALC. VALVAS
Ausente Leve Moderado Severo
ENGR. VALVAS
Normal Reduc. leve Reduc. mod. Reduc. sev.
MOVILIDAD
SCORE MITRAL: 9
Prolapso
Válvula mixoide
Engr. anillo
Ausente Leve Moderada Severa
CALC. COMIS.
Abiertas Fus. leve Fus. mod. Fus. severa
FUS. COMIS.
Simétrica Pred. ant. Pred. post.
Ausente Leve Moderada Severa
CALC. ANILLO
SAM
No vegetaciones
Sin afectación Afect. leve Afect. moderada Afect. severa
APARATO SUBVALV.
Elongación cuerdas tendíneas
Rotura cuerdas tendíneas
Rotura músculo papilar
Ausente Leve Moderada Severa
DILAT. ANILLO
Anterior
Resto normal
Continuar


INTRODUCIR ECO
Archivo Datos previos Hallazgos eco Diagnóstico Especial Ayuda
DIAGNOSTICO
Estenosis mitral reumática moderada. (100%)
Insuficiencia mitral leve. (96%)
Estenosis reumática severa de la válvula aórtica. (100%)
Insuficiencia tricuspídea funcional leve. (72%)
Retraso de la relajación diastólica. (65%)
Hipertensión pulmonar moderada. (100%)
Anterior
Resto normal
Continuar


INTRODUCIR ECO
Archivo Datos previos Hallazgos eco Diagnóstico Especial Ayuda
MEDICO
Dr. Río Aguilar
Dr. Enrique González
Dra. Elena Iturralde
Dr. Javier Jiménez
Dr. Iñigo Lozano
Dr. Javier Lozano
Dra. Sonia Rodríguez
Desconocido
Modificar
Anterior
Continuar

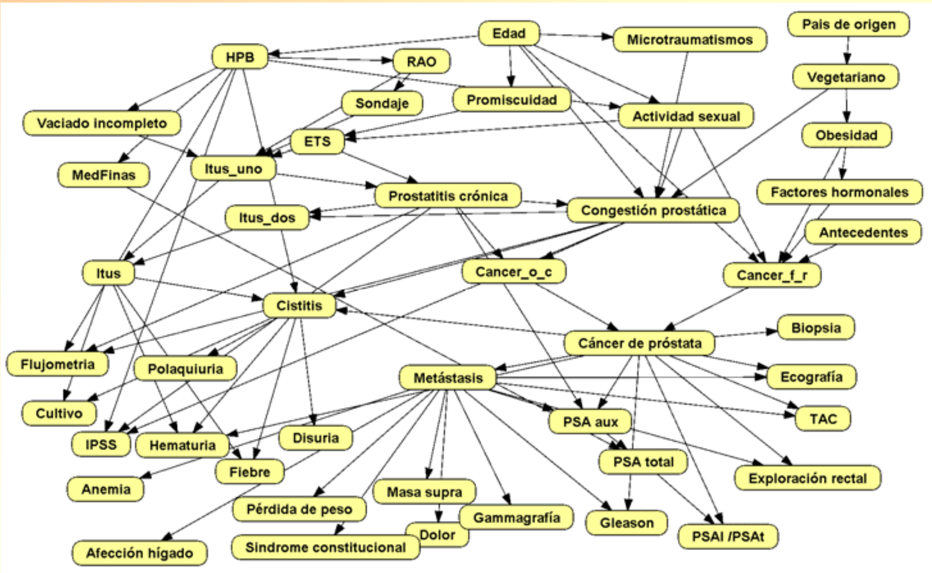


INFORME.TXT - Bloc de notas

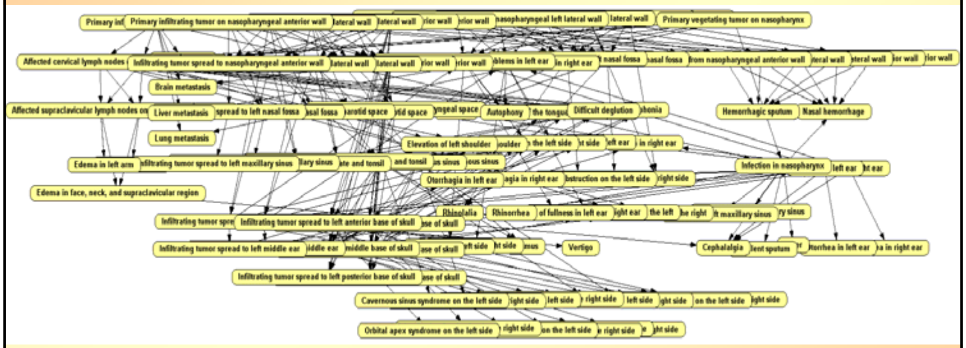
Archivo Edición Formato Ver Ayuda

```
Datos administrativos
======================
  Nº eco: 104382.  Fecha: 29/10/3.  Cinta: 512.
  MARIA GARCIA PEREZ.  DNI: 1234567.
  Edad: 51 años.  Mujer.
  Peso: 58 Kg.  Estatura: 158 cm.  Sup. corporal: 1.58 m².
  Solicitante: CARDIOLOGIA.
  Ingresada, sector 3, cama 512A.

Síntomas
=========
  Disnea de grado II.

Válvula mitral
===============
  Area (eco 2D): 1.2 cm².
  Velocidad onda E: 164 cm/s.
  Gradiente máximo: 10.8 mmHg.
  Gradiente medio: 7.0 mmHg.
  Tiempo de hemipresión: 255 ms.
  Area (THP): 0.9 cm².
  Engrosamiento moderado de las valvas.
  No calcificación de las valvas mitrales.
  Reducción leve de la movilidad.


[...]


Contractilidad segmentaria normal.

Pericardio normal.

DIAGNOSTICO
===========
  Estenosis mitral reumática moderada.
  Insuficiencia mitral leve.
  Estenosis reumática severa de la válvula aórtica.
  Insuficiencia tricuspídea funcional leve.
  Retraso de la relajación diastólica.
  Hipertensión pulmonar moderada.


                  Dra. Elena Iturralde  |
```

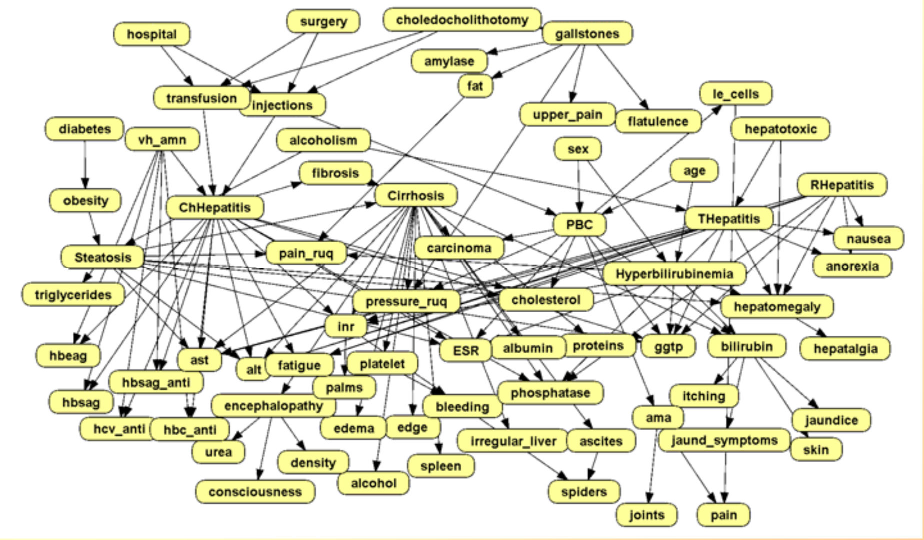
Prostanet (prostate diseases)



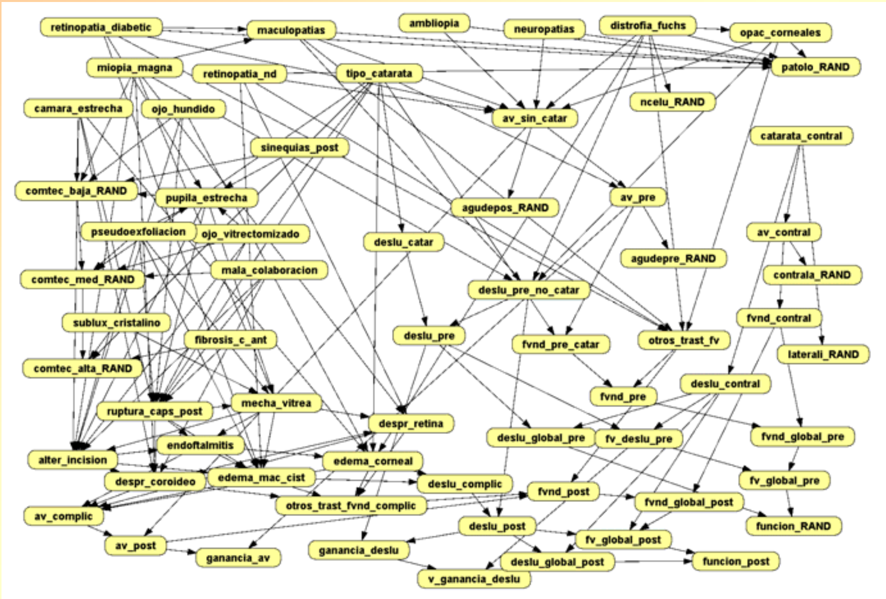
Nasonet (nasopharyngeal cancer)



Hepar II (liver diseases)



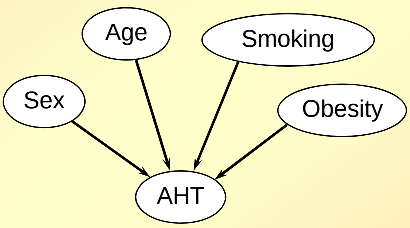
Catarnet (cataract surgery)



Canonical models

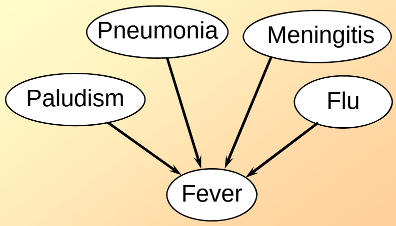
General model

- Probability table:
  $P(y \mid x_1, \ldots, x_n)$
- Factors that influence the prob. of X



Noisy OR

- Efficiency of each link:
  c_i
- Causes that may produce X



Our research on Bayesian networks (2/4)

- Canonical models
  - The noisy MAX, noisy AND and noisy MIN.
    - Díez. Parameter adjustment in BN's. The generalized noisy OR-gate. *UAI*, 1993
  - Inference with canonical models
    - Díez, Galán. Efficient computation for the noisy-MAX. 2003
  - A review of canonical models
    - Díez, Druzdzel. Canonical probabilistic models for knowledge engineering. 2007

Technical Report CISIAD-06-01 Version 0.9 (April 28, 2007)

Canonical Probabilistic Models for Knowledge Engineering

Francisco J. Díez FJDIEZ@DIA.UNED.ES
Dept. Inteligencia Artificial, UNED
Juan del Rosal, 16, 28040 Madrid, Spain

Marek J. Druzdzel MAREK@SIS.PITT.EDU
Decision Systems Laboratory, School of Information Sciences and Intelligent Systems Program
University of Pittsburgh, Pittsburgh, PA 15260, USA

Abstract

The hardest task in knowledge engineering for probabilistic graphical models, such as Bayesian networks and influence diagrams, is obtaining their numerical parameters. Models based on acyclic directed graphs and composed of discrete variables, currently most common in practice, require for every variable a number of parameters that is exponential in the number of its parents in the graph, which makes elicitation from experts or learning from databases a daunting task. In this paper, we review the so called *canonical models*, whose main advantage is that they require much fewer parameters. We propose a general framework for them, based on three categories: deterministic models, ICI models, and simple canonical models. ICI models rely on the concept of *independence of causal influence* and can be subdivided into noisy and leaky. We then analyze the most common families of canonical models (the OR/MAX, the AND/MIN, and the noisy XOR), generalizing them and offering criteria for applying them in practice. We also briefly review temporal canonical models.

Contents

Our research on Bayesian networks (3/4)

- Explanation in BNs
  - Review of the literature
    - Lacave, Díez. A review of explanation methods for Bayesian networks. 2002.
  - New explanation facilities, implemented in Elvira
    - Lacave, Luque, Díez. Explanation of BNs and IDs in Elvira. 2007.
  - which are useful for tuition
    - Díez. Teaching probabilistic medical reasoning with the Elvira software. 2004
  - and for building and debugging BNs
    - Lacave, Onisko, Díez. Use of Elvira's explanation facility for debugging probabilistic expert systems. 2006.

The Knowledge Engineering Review, Vol. 17:2, 107–127. © 2002, Cambridge University Press
DOI: 10.1017/S026988890200019X Printed in the United Kingdom

A review of explanation methods for Bayesian networks*

CARMEN LACAVE[1] and FRANCISCO J DÍEZ[2]

[1] *Dept. Computer Science, University of Castilla-La Mancha, Paseo de la Universidad, s/n 13071 Ciudad Real, Spain; e-mail: clacave@inf-cr.uclm.es*
[2] *Dept. Artificial Intelligence, UNED, Senda del Rey, 9, 28040 Madrid, Spain; e-mail: fjdiez@dia.uned.es*

Abstract

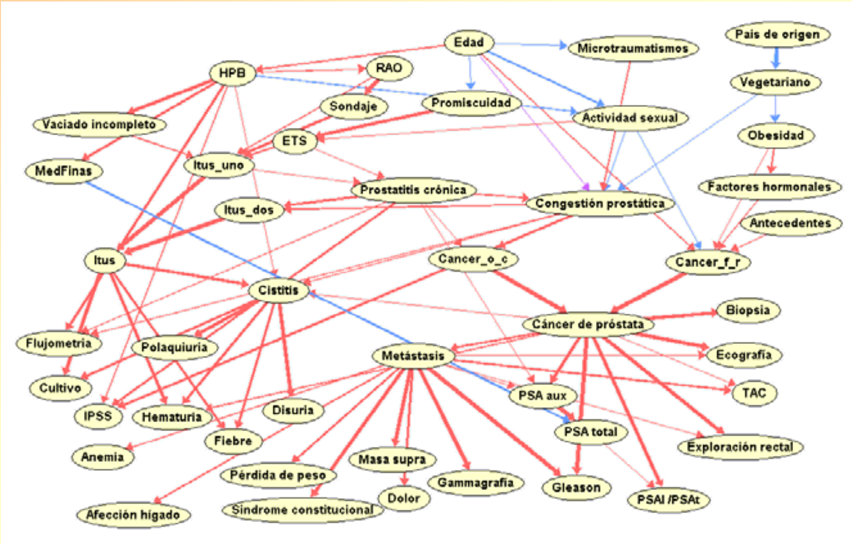
One of the key factors for the acceptance of expert systems in real-world domains is the ability to explain their reasoning (Buchanan & Shortliffe, 1984; Henrion & Druzdzel, 1990). This paper describes the basic properties that characterise explanation methods and reviews the methods developed to date for explanation in Bayesian networks.

1 Introduction

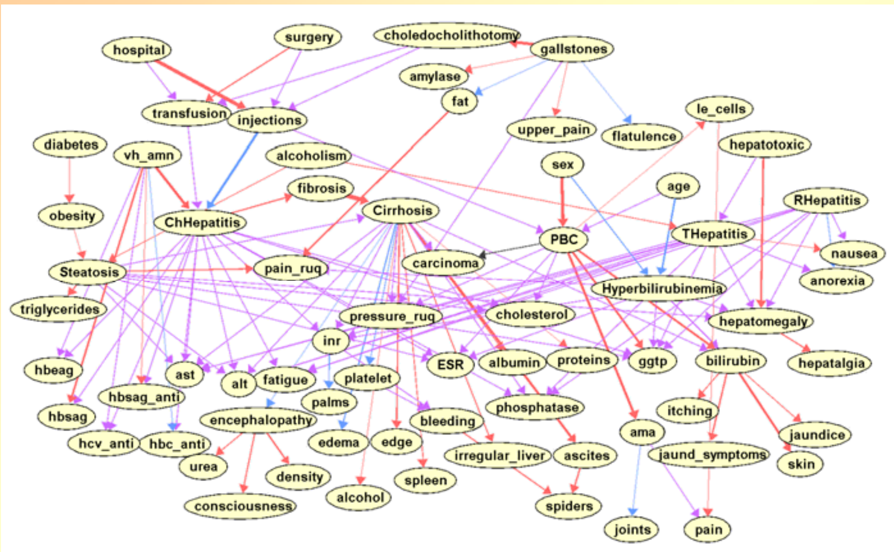
Expert systems originated in the 1970s as computer programs capable of imitating human experts and even substituting them when necessary. One of the essential qualities of real experts is their ability to communicate their knowledge and explain their reasoning. This ability is especially important in the case of expert systems, not only for tracing performance during the construction and evaluation of the system, but also for justifying their results when the system is deployed in an operating environment. In fact, an experiment performed at the MYCIN project showed that physicians are very reluctant to accept the advice of a machine if they do not understand how it was obtained (Teach & Shortliffe, 1984).

In the decades that followed, i.e. the 1980s and 1990s, the main goal of artificial intelligence shifted from *imitating* natural intelligence to *supporting* human beings in a synergistic way. In fact, Clancey (1993) points to the notes to authors' in the *Knowledge Acquisition* journal: "The key issue is not

Sign of influences (coloring of links)



Sign of influences (coloring of links)



How to build a Bayesian network

- From a database

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

  - There are many algorithms, several new algorithms every year
  - Similar to statistical methods (logistic regression, neural nets...)

- With a human expert's help

Causal knowledge $\xrightarrow{\text{modeling}}$ Causal graph $\xrightarrow{\text{probabilities}}$ Bayesian network

- Hybrid methods:
  - experts → structure; database → probabilities
  - experts → initial model; new cases → refine the probabilities

Our research on Bayesian networks (4/4)

- Learning Bayesian networks interactively
  - The system proposes, the user decides
  - Very useful for tuition
  - Useful for combining data with expert knowledge
  - Useful for debugging new algorithms (workbench).
  - Implemented in OpenMarkov:
    www.openmarkov.org/docs/tutorial

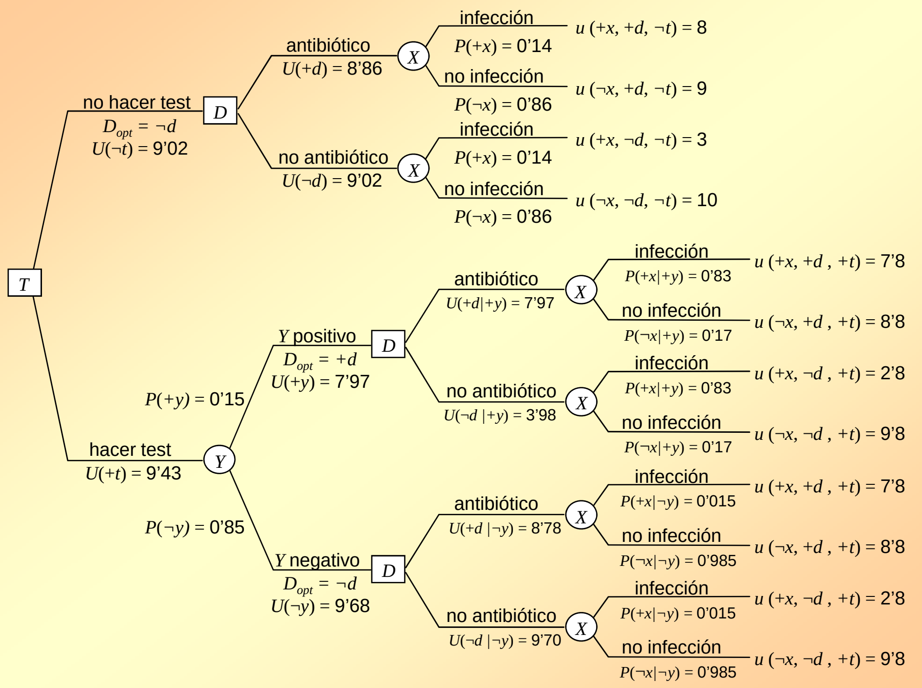
Influence diagrams

A medical problem

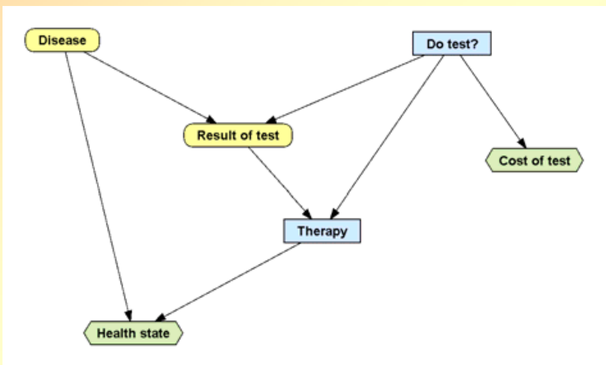
- Disease X
  - Prevalence: $P(+x) = 0'14$
- Therapy D
  - Utility:

$u(x, d)$	$+x$	$\neg x$
$+d$	8	9
$\neg d$	3	10

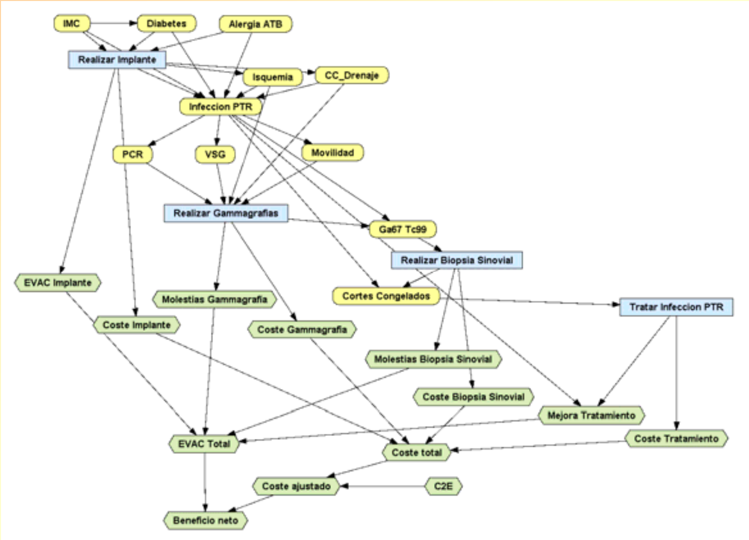
- Test Y
  - Sensitivity: $P(+y|+x) = 0'91$
  - Specificity: $P(\neg y|\neg x) = 0'97$
  - Cost: $u_{\text{test}}(x, d) = u_{\text{not-test}}(x, d) - 0'2$
- Decisions:
  - Is it worth to do the test?
  - In what cases should we apply the therapy?



An ID for this example

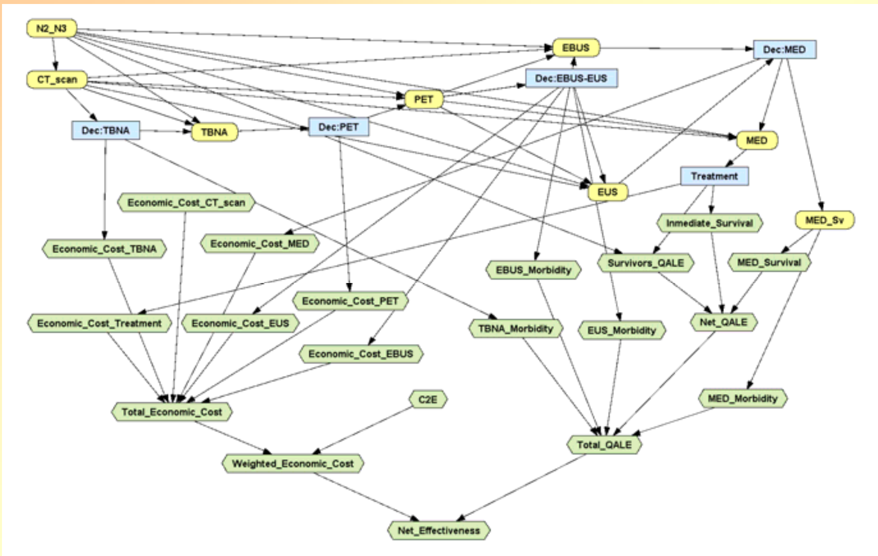


Arthronet (total knee arthroplasty)



Equivalent to a decision tree containing ~10^4 branches.

Mediastinet (lung cancer)



Equivalent to a decision tree containing ~10^7 branches.

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 (mortatility, morbidity, economic cost...)
  - ID can use canonical models (noisy OR, noisy AND, etc)
  - Each parameter appears only once in the ID
    - in many cases it is not necessary to have parametric variables
  - IDs can use super-value nodes: explicit combination of utilities

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*)
- No external pre-calculation of probabilities is required
- IDs are much easier to modify than decision trees
  - Refine the model with new decisions and chance variables
  - Structural sensitivity analysis
  - 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.

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

DECISION ANALYSIS
Vol. 2, No. 4, December 2005, pp. 229–231
ISSN 1545-8490 | EISSN 1545-8504 | 05 | 0204 | 0229

informs.

DOI 10.1287/deca.1050.0054

The Influence of Influence Diagrams on Artificial Intelligence

Craig Boutilier
Department of Computer Science, University of Toronto, Toronto, Ontario, M5S 3G5 Canada, cebly@cs.toronto.edu

Howard and Matheson's article "Influence Diagrams" has had a substantial impact on research in artificial intelligence (AI). In this perspective, I briefly discuss the importance of influence diagrams as a model for decision making under uncertainty in the AI research community; but I also identify some of the less direct, but no less important, influences this work has had on the field.

Key words: influence diagrams; decision theory; artificial intelligence; value of information; graphical models; perspective, the focus on graphical modeling research
History: Received on November 14, 2005. Accepted by Eric Horvitz on November 23, 2005, without revision.

Howard and Matheson's (1984/2005) "Influence Diagrams" has had a profound impact on developments in artificial intelligence. Some of these influences have been quite direct; others are more indirect, but in many ways, more substantial. The paper vision (Binford and Levitt 2003), dialog management, user interface design, multiagent systems, and game theory (Koller and Milch 2003), to name but a few.

Another reasonably direct impact of "Influence Diagrams" derives from its role in the development

DECISION ANALYSIS
Vol. 2, No. 4, December 2005, pp. 238–244
ISSN 1545-8490 | EISSN 1545-8504 | 05 | 0204 | 0238

informs
DOI 10.1287/deca.1060.0060

The Influence of Influence Diagrams in Medicine

Stephen G. Pauker, John B. Wong
Division of Clinical Decision Making, Informatics and Telemedicine, Department of Medicine,
Tufts–New England Medical Center, Tufts University School of Medicine, 750 Washington St., NEMC 302,
Boston, Massachusetts 02111 {spauker@tufts-nemc.org, jwong@tufts-nemc.org}

Although influence diagrams have used medical examples almost from their inception, that graphical representation of decision problems has disseminated surprisingly slowly in the medical literature and among clinicians performing decision analyses. Clinicians appear to prefer decision trees as their primary modeling metaphor. This perspective examines the use of influence diagrams in medicine and offers explanations and suggestions for accelerating their dissemination.

Key words: decision analysis; influence diagrams; clinical decision making; medicine
History: Received December 12, 2005. Accepted by Eric Horvitz on January 5, 2006, after 1 revision.

Introduction

Two decades after Howard's landmark paper (Howard and Matheson 1984/2005) that introduced the concept of the influence diagram and three decades since Miller's initial report (Miller et al. 1976), *Decision Analysis* reproduced that paper in 2005 and solicited a set of commentaries. This paper

modeling paradigm slowly spread from Stanford, both with courses offered at meetings of the Society for Medical Decision Making (Society for Medical Decision Making 2005) and with the development of software that could conveniently capture and evaluate such models.

IDs in the literature on MDM (1/3)

- Books that mention decision trees but do not mention IDs
  - Weinstein, Fineberg. *Clinical Decision Making*. 1980.
  - Sloan (ed.). *Valuing Health Care*. 1995.
  - Gold et al. *Cost-Effectiveness in Health and Medicine*. 1996.
  - Sackett et al. *Evidence-Based Medicine*. 1997 (and three other books on EBM).
  - Petiti. *Meta-Analysis, Decision Analysis and CEA*. 2nd ed., 2000.
  - Drummond, McGuire (eds.). *Economic Eval. in Health Care Programs*. 2001.
  - Levin and McEwan. *Cost-Effectiveness Analysis*. 2nd ed., 2001.
  - Parmigiani. *Modelling in Medical Decision* Making. 2002.
  - Haddix et al. *Prevention Effectiveness*. 2nd ed., 2003.
  - Briggs et al. *Decision Modelling for Health Economic Evaluation*, 2006.
  - Kassirer et al. *Learning Clinical Reasoning*. 2nd ed., 2010.
  - Mushlin and Greene. *Decision Making in Medicine*. 3rd ed., 2010.
  - Gray et al. *Applied Methods of CEA in Health Care*, 2011.

(cont'd)

IDs in the literature on MDM (2/3)

- Books that mention decision trees but do not mention IDs (cont.)
  - Alfaro-LeFevre. *Critical Thinking, Clinical Reasoning, and Clinical Judgment*. 5th ed., 2013.
  - Sox et al. *Medical Decision Making*. Latest edition: 2013.
  - Hunink et al. *Decision Making in Health and Medicine*. 2nd ed., 2014.
  - Drummond et al. *Methods for the Economic Evaluation of Health Care Programmes*. 4th ed. 2015.
- Two books that mention IDs
  - Chapman and Sonnenberg (eds.). *Decision Making in Health Care*. 2000 (5 pages out of 421, in a chapter authored by Mark Roberts).
  - Schwartz and Bergus. *Medical Decision Making. A Physician's Guide*. 2008.
- Another book that mentions IDs
  - Muenning. *Designing and Conducting Cost-Effectiveness Analyses in Medicine and Health Care*. 2002.

    "An influence diagram (also known as a tornado diagram) ..." [p. 242]

    The mistake is (partially) corrected in the second edition of the book, 2008.

IDs in the literature on MDM (3/3)

- Summary of the informal survey of books on MDM and EBM
  - 22 books published after 1984
  - All of them explain DTs but only two describe IDs, very briefly.
- Some books on medical informatics that mention IDs:
  - Shortliffe and Cimino. *Biomedical Informatics*. 4th ed., 2013 (2.5 pages out of 991).
  - Kalet. *Principles of Biomedical Informatics*. 2nd ed., 2013 (3 pages out of 708).
- Why are IDs almost unknown in health sciences after 30+ years?

Limitations of IDs

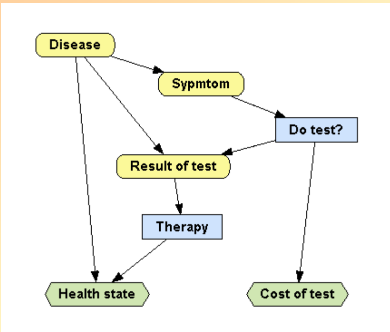
1. The “reasoning” of an ID is not easy to understand
2. The evaluation returns large policy tables
3. Algorithms could only evaluate unicriterion IDs
   - They cannot perform cost-effectiveness analysis
4. Temporal reasoning was not possible with IDs
   - Dynamic IDs are computationally unfeasible.
5. IDs cannot model symmetric problems
   - IDs require a total ordering of the decisions
   - IDs cannot represent incompatibilities between values
     - Non-standard versions of IDs partially solve this problem, but none of the alternatives is completely satisfactory.

Solutions we have proposed

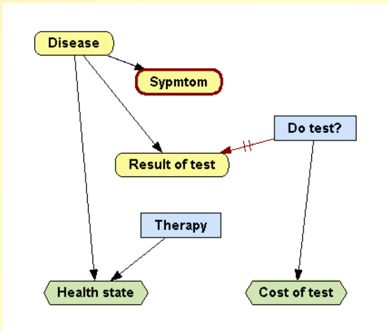
1. Explanation in influence diagrams
   - showing the posterior probabilities and expected values
   - introduction of evidence
   - hypothetical reasoning (what if) by means of imposed policies
2. Synthesizing the optimal intervention
   - in the form of a compact tree
3. Cost-effectiveness analysis with IDs
4. Markov influence diagrams
   - including cost-effectiveness analysis
5. Decision analysis networks
   - an alternative to IDs for asymmetric decision problems.

Decision analysis networks

Influence diagram



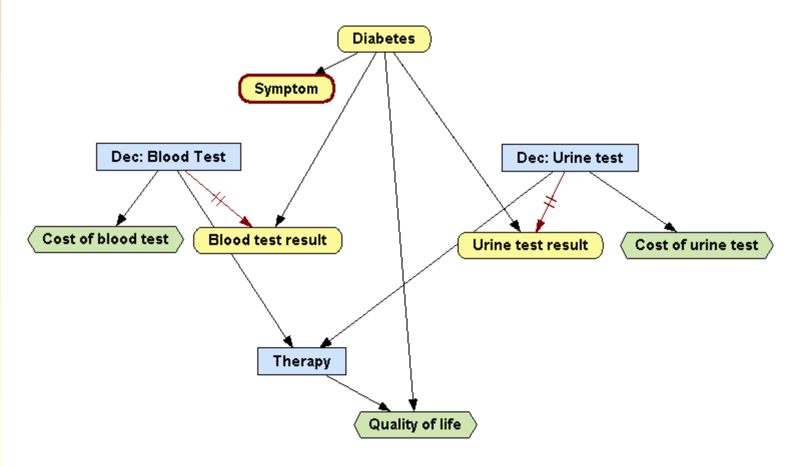
Decision analysis network



- The ID contains two information arcs:
  - because the symptom is always observed (spontaneously)
  - because the result of the test is known just after doing the test
- The variable “Result of test” does not make sense when the test is not performed

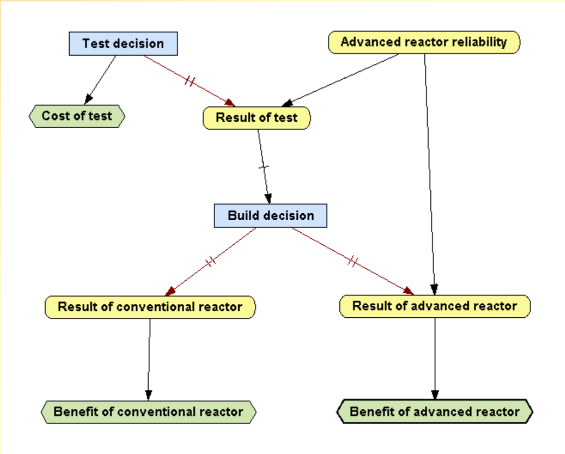
The diabetes problem

[Demirer & Shenoy, 2001]



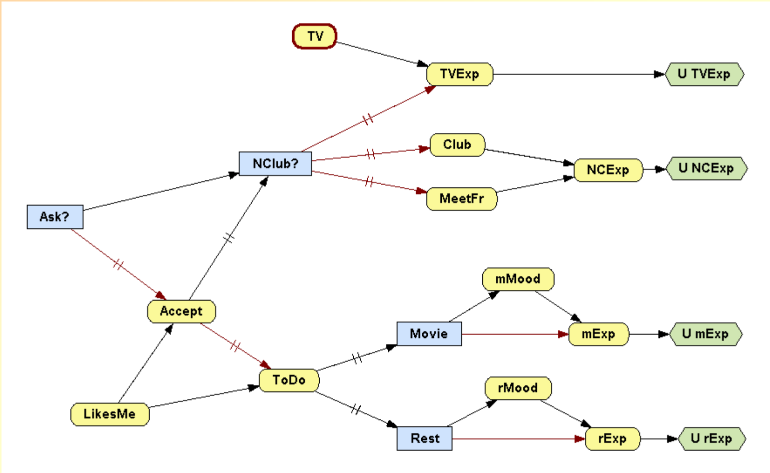
The reactor problem

[Covaliu and Oliver, 1995; Bielza and Shenoy, 1999]

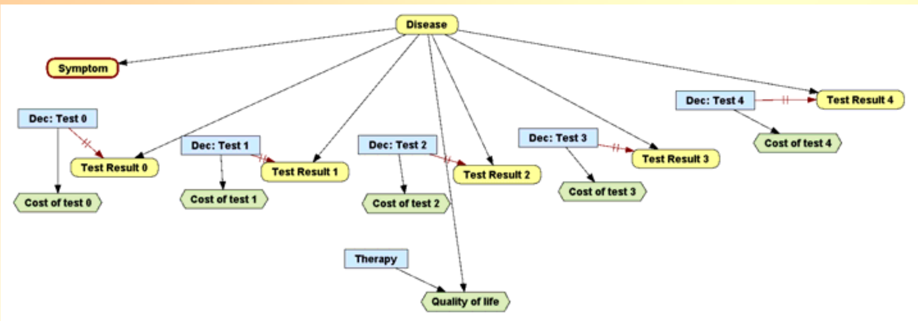


The dating problem

[Nielsen & Jensen, 2000; Jensen et al., 2004]

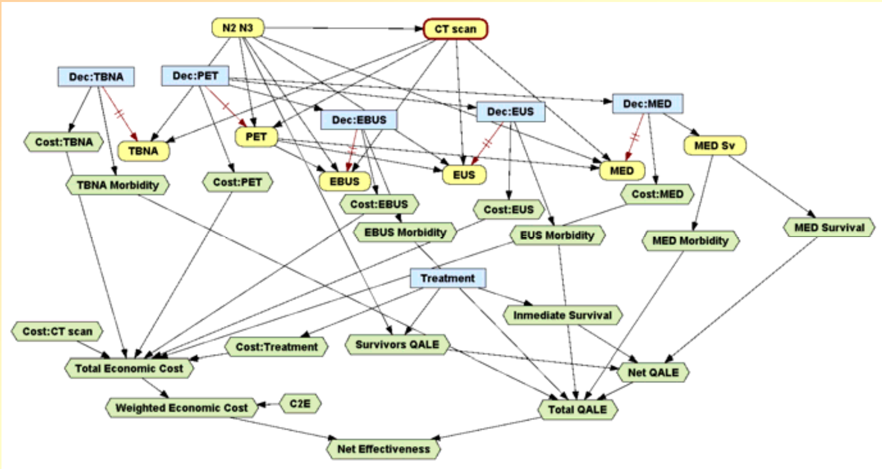


The *n*-test problem



- Computationally complex: $n!$ possible orderings of the tests.
- We have developed an any-space algorithm for this problem
- and a fast algorithm (9 minutes for the 7-test problem).
- We are developing more efficient algorithms.

Mediastinet (DAN version)



Decisions are partially ordered.

DANs vs. IDs

- DANs can replace IDs as the standard decision analysis tool (in AI, MDM, operations research…) for these reasons:
  - For every ID there is an equivalent symmetric DAN
  - Virtually all real-world problems are asymmetric.
  - There many problems that cannot be modeled with IDs.
  - Even if a problem can be modeled with an ID, a DAN is usually better because it does not need dummy states.

Cost-effectiveness analysis with influence diagrams

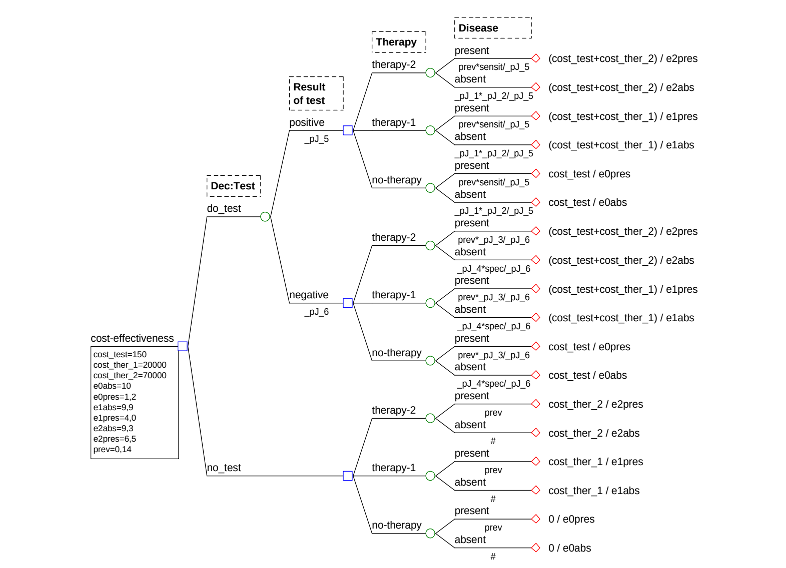
Example: Cost-effectiveness of a test

- Disease prevalence = 0.14
- Test sensitivity = 0.90, specificity = 0.93, cost = 150 €
- Therapy 1 cost = 20,000 €
- Therapy 2 cost = 70,000 €
- Effectiveness (QALYs)

	No therapy	Therapy 1	Therapy 2
Disease present	1,2	4,0	6,5
Disease absent	10	9,9	9,3

- Is the test cost-effective?
- What is the most cost-effective therapy?

A decision tree for cost-effectiveness analsysis



TreeAge usually returns wrong results for this tree.

PharmacoEconomics (2014) 32:1141–1145
DOI 10.1007/s40273-014-0195-1

RESEARCH LETTER

The Problem of Embedded Decision Nodes in Cost-Effectiveness Decision Trees

Manuel Arias · Francisco Javier Díez

Published online: 31 July 2014

1 Introduction

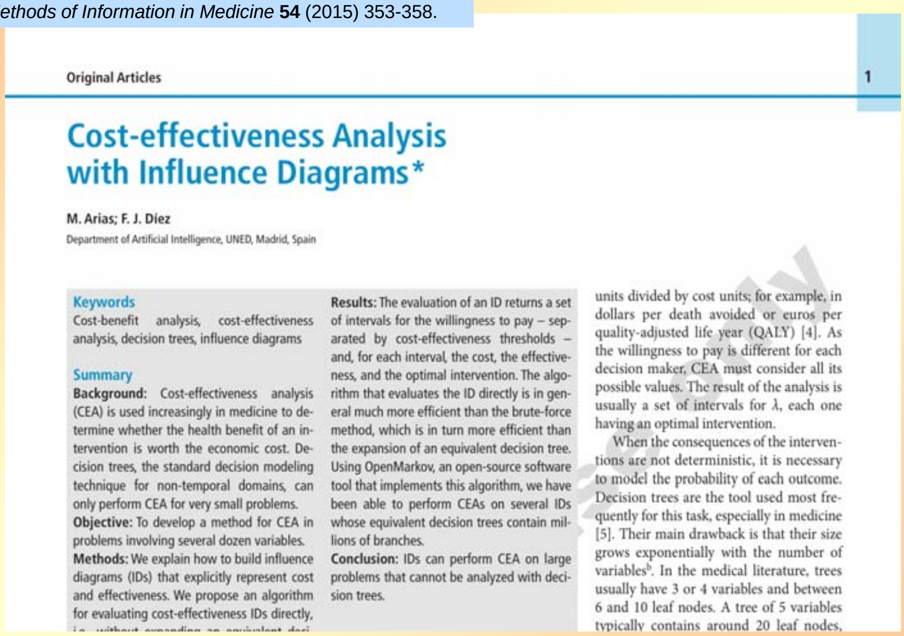
Cost-effectiveness analysis (CEA) is increasingly used to inform health policies. Decision trees are the standard method for decision analysis in non-temporal domains. A decision node that is not the root of the tree is said to be embedded.

All books on medical decision analysis discuss both CEA *and* decision trees [1–11], but few explain how to conduct a CEA *with* decision trees [1, 2, 10, 11], and only build a decision tree with one decision node at the root, which represents all the strategies to be evaluated, as proposed by Kuntz and Weinstein; the other is to apply the algorithm presented in Arias and Díez [13].

As a case study, we consider the common problem of finding the incremental cost-effectiveness ratio (ICER) of a test:

Example 1 For a disease with a prevalence of 0.14, there are two possible therapies, the effectiveness of which depends on whether or not the disease is present, as shown

Methods of Information in Medicine **54** (2015) 353-358.



Original Articles 1

Cost-effectiveness Analysis with Influence Diagrams*

M. Arias; F. J. Díez
Department of Artificial Intelligence, UNED, Madrid, Spain

Keywords
Cost-benefit analysis, cost-effectiveness analysis, decision trees, influence diagrams

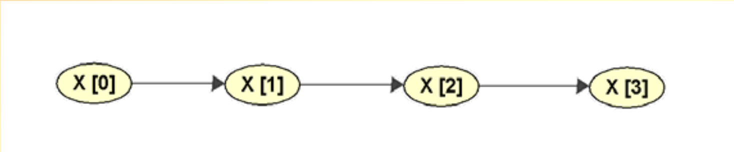
Summary
Background: Cost-effectiveness analysis (CEA) is used increasingly in medicine to determine whether the health benefit of an intervention is worth the economic cost. Decision trees, the standard decision modeling technique for non-temporal domains, can only perform CEA for very small problems.
Objective: To develop a method for CEA in problems involving several dozen variables.
Methods: We explain how to build influence diagrams (IDs) that explicitly represent cost and effectiveness. We propose an algorithm for evaluating cost-effectiveness IDs directly,
Results: The evaluation of an ID returns a set of intervals for the willingness to pay – separated by cost-effectiveness thresholds – and, for each interval, the cost, the effectiveness, and the optimal intervention. The algorithm that evaluates the ID directly is in general much more efficient than the brute-force method, which is in turn more efficient than the expansion of an equivalent decision tree. Using OpenMarkov, an open-source software tool that implements this algorithm, we have been able to perform CEAs on several IDs whose equivalent decision trees contain millions of branches.
Conclusion: IDs can perform CEA on large problems that cannot be analyzed with decision trees.

units divided by cost units; for example, in dollars per death avoided or euros per quality-adjusted life year (QALY) [4]. As the willingness to pay is different for each decision maker, CEA must consider all its possible values. The result of the analysis is usually a set of intervals for λ, each one having an optimal intervention.

When the consequences of the interventions are not deterministic, it is necessary to model the probability of each outcome. Decision trees are the tool used most frequently for this task, especially in medicine [5]. Their main drawback is that their size grows exponentially with the number of variables[b]. In the medical literature, trees usually have 3 or 4 variables and between 6 and 10 leaf nodes. A tree of 5 variables typically contains around 20 leaf nodes,

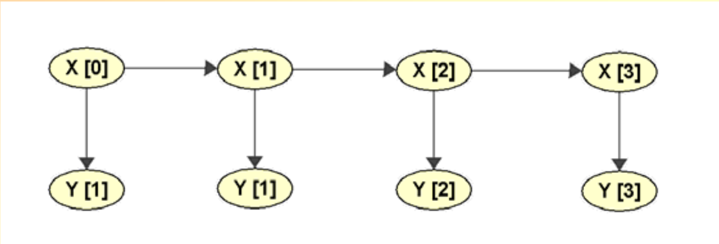
PGMs for temporal reasoning

Markov chain



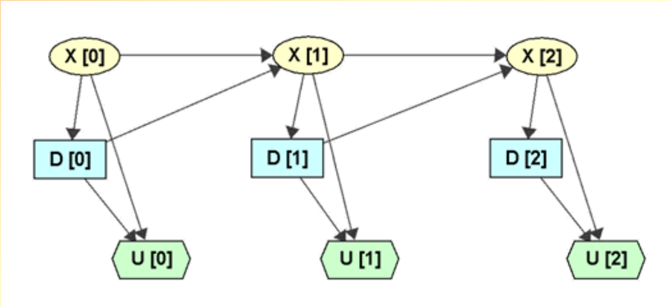
- One variable that evolves over time
- Transition probabilities: $P(x_{i+1}|x_i)$

Hidden Markov model (HMM)



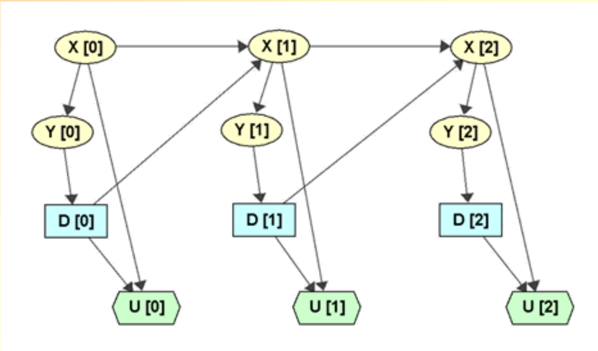
- Observed variable: Y
- Non-observed (hidden) variable: X
- Probability of each observation: $P(y_i|x_i)$
- Transition probability: $P(x_{i+1}|x_i)$

Markov decision process (MDP)



- Observed variable: X
- Decision: D
- Transition probability: $P(x_{i+1}|x_i)$
- Reward: $U(x_i, d_i)$

Partially observable MDP (POMDP)

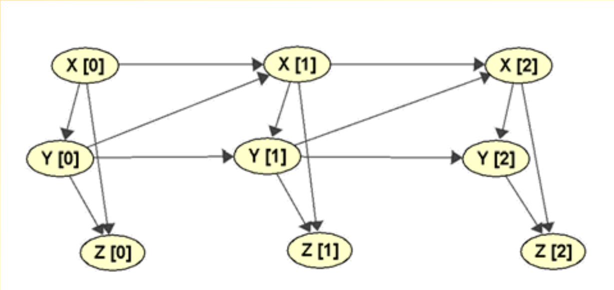


- Hidden variable: X
- Observed variable : Y
- Decision: D
- Observation prob.: $P(y_i|x_i)$
- Transition prob.: $P(x_{i+1}|x_i)$
- Reward: $U(x_i, d_i)$

Factored extensions of Markov models

Flat model	Factored model
Markov chain	Dynamic Bayesian network [Dean and Kanazawa, 1989]
Hidden Markov model	
Markov decision process (MDP)	Factored MDP [Boutilier et al., 1995, 2000]
Partially-observable MDP (POMDP)	Factored POMDP [Boutilier and Poole, 1996]

Dynamic Bayesian network (DBN)



- Markov chain or hidden Markov model:
  - one variable, X
  - one conditional probability: $P(x_{i+1}|x_i)$
- Dynamic Bayesian network:
  - several variables, $\{X, Y, Z\ldots\}$
  - factored probability: $P(y_i|x_i)$, $P(z_i|x_i, y_i)$, $P(x_{i+1}|x_i, y_i)\ldots$

IJCAI Workshop *Decision Making in Partially Observable, Uncertain Worlds: Exploring Insights from Multiple Communities*
Barcelona, July 2011

MDPs in Medicine: Opportunities and Challenges

F. J. Díez M. A. Palacios M. Arias
Dept. Artificial Intelligence. UNED
Madrid, Spain

Abstract

In the last three decades hundreds of Markov models have been built for medical applications, but most of them fall under the paradigm of what we call *simple Markov models* (SMMs). Markov decision processes (MDPs) are much more powerful as a decision analysis tool, but they are ignored in medical decision analysis books and the number of medical applications based on them is still very small. In this paper we compare both types of models and discuss the challenges that MDPs must overcome before they can be widely accepted in medicine. We present a software tool, OpenMarkov, that addresses those challenges and has been used to build a Markov model for analyzing the cost-effectiveness of the HPV vaccine.

1 Introduction

Markov models were introduced in the beginning of the 20-th century by the Russian mathematician Andrei Andreyevich Markov [1906]. In the three decades passed since the pioneering work of Beck and Pauker [1983], hundreds of

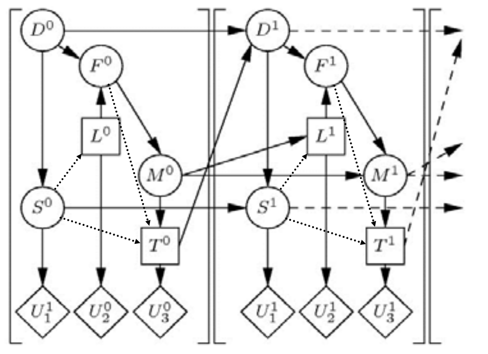
the emergence of partially observable Markov decision processes (POMDPs) [Åström, 1965], in which the state of the system is not directly observable, but there is a variable that correlates probabilistically with it. POMDPs were developed in the field of automatic control as an extension of MDPs, but currently most of the research about them is carried out in artificial intelligence (AI), again as a tool for planning, especially in robotics [Ghallab *et al.*, 2004]. The main contribution of AI to this field comes from the area of probabilistic graphical models: Bayesian networks [Pearl, 1988] led to the development of dynamic Bayesian networks [Dean and Kanazawa, 1989], which generalize Markov chains and hidden Markov models [Murphy, 2002]. The idea of using several variables to represent the state of the system, instead of only one, led to factored MDPs [Boutilier *et al.*, 1995; 2000] and factored POMDPs [Boutilier and Poole, 1996], which can model efficiently many problems that were unmanageable with flat (i.e., non-factored) representations; correspondingly, there are new algorithms that can solve problems several orders of magnitude bigger than in the recent past [Hoey *et al.*, 1999; Poupart, 2005; Spaan and Vlassis, 2005].

In the rest of the paper, we use the acronym MDPs to refer to both fully observable and partially observable models

Our research on temporal PGMs

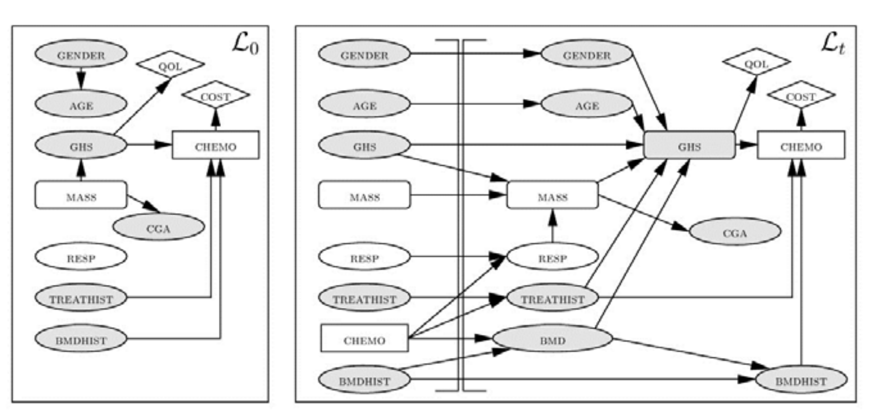
- New types of models
  - Networks of probabilistic events in discrete time (NPEDT)
    - A non-Markovian extension of BNs
  - Dynamic limited-memory influence diagrams (LIMID)
    - A Markovian extension of limited-memory IDs [Lauritzen & Nilsson, 2001]
  - Markov influence diagrams (MID)
    - A Markovian extension of influence diagrams
- An algorithm for cost-effectiveness analysis with MIDs
  - and several MIDs for different medical problems.

Dynamic limited-memory IDs (DLIMIDs)



- Differences from POMDPs
  - Several decisions in each time slice.
  - Limited memory: the decision maker only knows the observations from the current and previous time slices
  - Memory variables summarize the past.

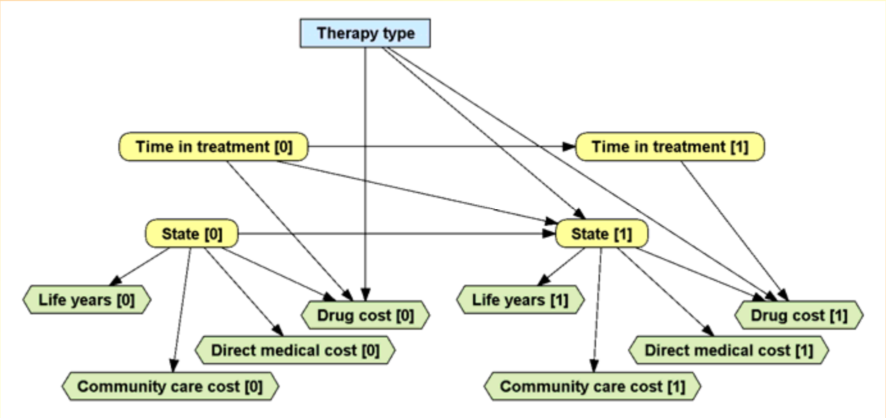
A DLIMID for a carcinoid tumors



- Therapy selection for high-grade carcinoid tumors (van Gerven et al., 2007)

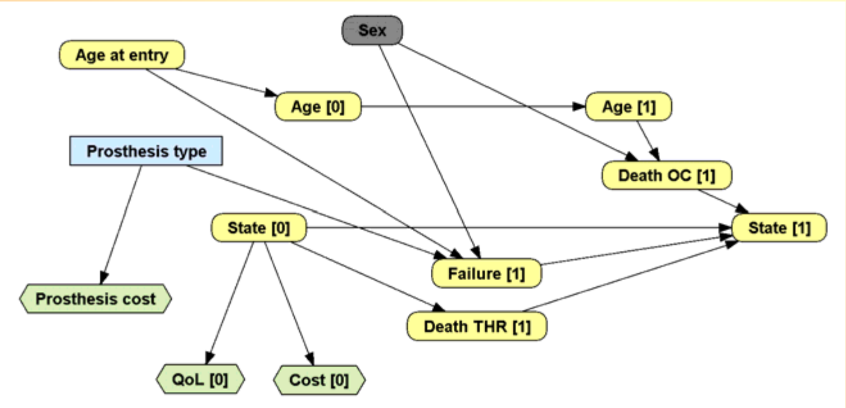
A MID version of the HIV model

[Chancellor et al., 1997]



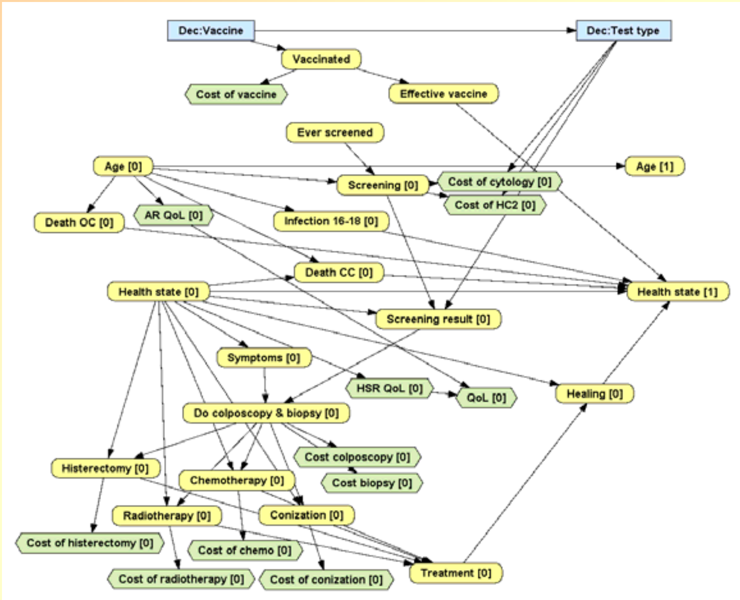
A MID version of the hip replacement model

[Briggs et al., 2004]



A MID version of the HPV vaccination model

[Callejo et al., 2010]

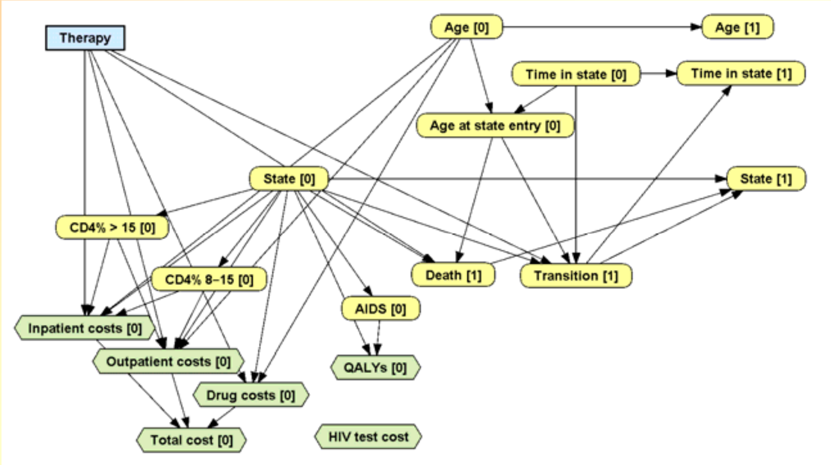


Content of one of the Excel cells for this model:

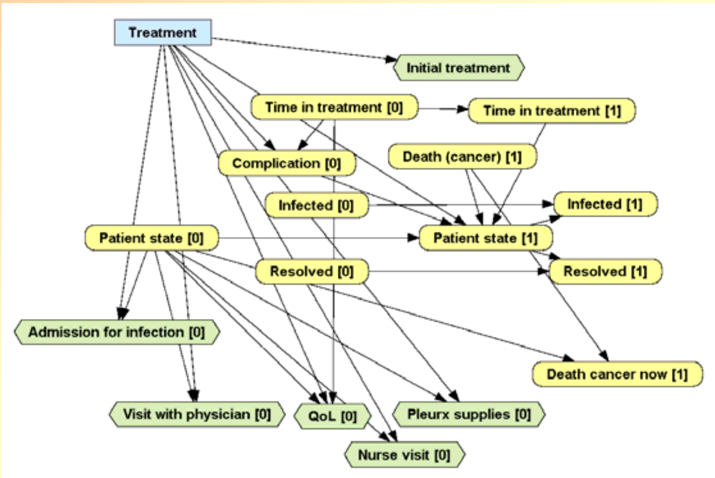
=VLOOKUP($C5;Variables!$A$4:$H$21;8;TRUE)*(((BI5+BJ5)+BK5*uCIN1+SUM(BL5:BP5)*uCIN2_3+(BQ5+BR5)*uLCC+(BS5+BT5)*uRCC+(BU5+BV5)*uDCC)+((BI4+BJ4)+BK4*uCIN1+SUM(BL4:BP4)*uCIN2_3+(BQ4+BR4)*uLCC+(BS4+BT4)*uRCC+(BU4+BV4)*uDCC)*VLOOKUP($C4;Variables!$A$4:$H$21;2;TRUE)+(BQ4+BR4)*uLCC*VLOOKUP($C4;Variables!$A$4:$H$21;4;TRUE)+(BS4+BT4)*uRCC*VLOOKUP($C4;Variables!$A$4:$H$21;5;TRUE)+(BU4+BV4)*uDCC*VLOOKUP($C4;Variables!$A$4:$H$21;2;TRUE))

A MID version of the CHAP model

[Ryan et al., 2008]

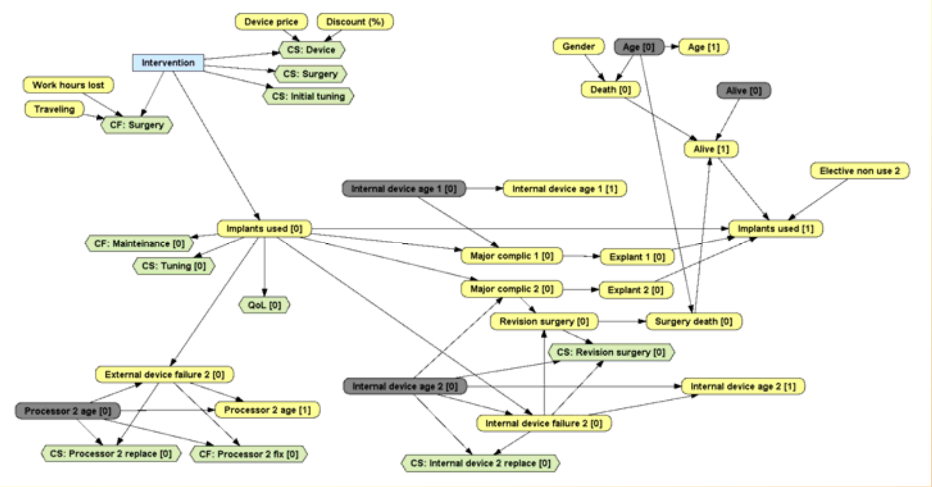


Our model for malignant pleural effusion



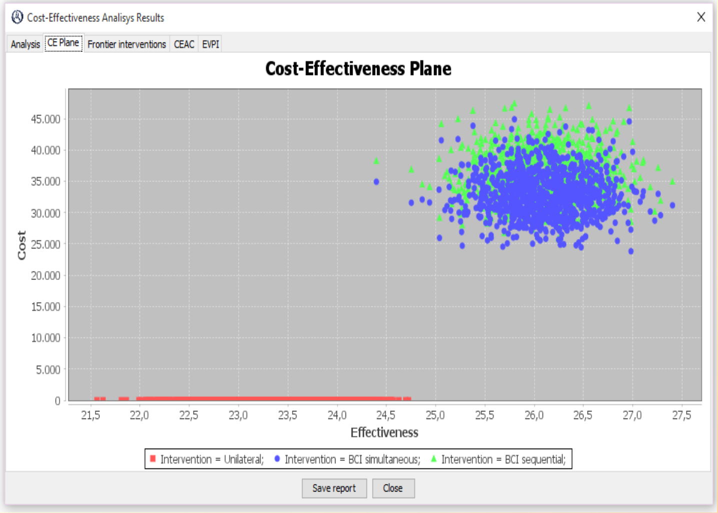
- *Meeting of the Society for Medical Decision Making*, St. Louis, October 19-21, 2015.

Our model for bilateral cochlear implantation

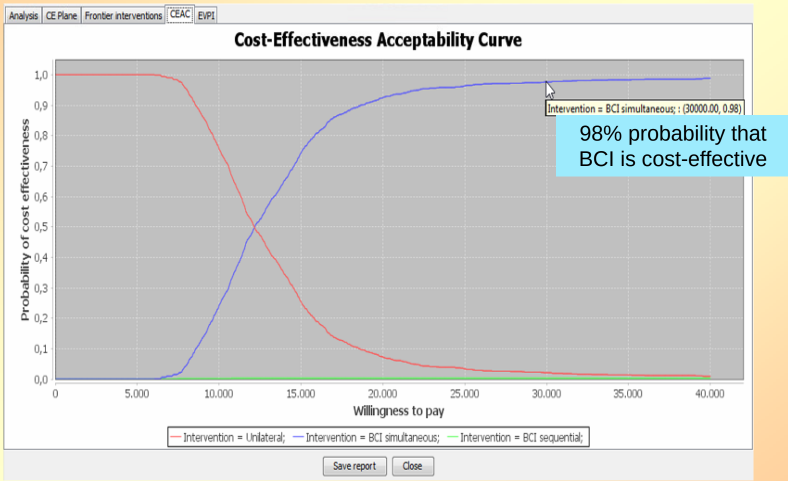


- *Cochlear Implant Symposium*, Washington DC, October 15-17, 2015.

Probabilistic sensitivity analysis



Probabilistic sensitivity analysis



Conclusions (1/2)

- In the first decades of AI probabilistic methods seemed to be inappropriate for reasoning and decision making.
- BNs overcame the limitations of the naïve Bayes method.
- IDs have several advantages over decision trees,
  but also have serious limitations for medical decision making.
- The main contributions of our group are:
  - new types of models: DANs, DLIMIDs, Markov IDs…;
  - new methods for the explanation of reasoning;
  - new algorithms, especially for cost-effectiveness analysis;
  - an open-source software package that implements these contributions (and the interactive learning of BNs); and
  - an XML format for encoding probabilistic models.

Conclusions (2/2)

- PGMs play a more and more relevant role in AI: robotics, planning, natural language, learning…
- They are the preferred approach in medicine:
  - normative basis: decision theory
  - combine expert knowledge with data
- What remains to be done:
  - dissemination in the fields of MDM and health economics
    - seminars, short courses, MOOC…
    - tutorials, journal papers, book…
  - better software tools:
    - sensitivity analysis, cost-effectiveness analysis, explanation…

Thank you very much for your attention!