Graphical Models for Decision Making

Lesson 3, week 13, class 25

Juan A Fdez del Pozo (CIG-UPM)

December 2, 2015

Contens

1 IdR

Outline of Contens

1 IdR

Graphical Models for Decision Making

IdR, intro

Influence Diagrams on R: IdR

Keywords: Machine Learning, Probabilistic Graphical Models, Bayesian Networks, Classification, Decision Making

- Decision making models: Bayesian Networks and Influence Diagrams
- Decision model evaluations
- Analysis and explanations of the results

** An educational package for Probabilistic Graphical Models
http://www.dia.fi.upm.es/~jafernan/research/idr/idr.html
http://www.dia.fi.upm.es/~jafernan/research/idr/IdR_1.0.zip

YARP: Yet Another R Package?

Educational and Research purposes about Probabilistic Graphical Models using the R environment (RGui, lattice, cluster, gat,...)

- Grappa: R functions for probability propagation,
  http://www.maths.bris.ac.uk/p̃eter/Grappa/
  Peter J. Green
  University of Bristol, UK
- Related R packages
  CRAN Task View: gRaphical Models in R
  deal, bnlearn, MASTINO,...
- Other Software (not in R):
  GeNIe & SMILE: http://genie.sis.pitt.edu (*)
  Hugin: http://www.hugin.com
  Elvira: http://www.leo.ugr.es/elvira/
  Ace, the Bayesian network compiler: http://reasoning.cs.ucla.edu/ace
  Mark Chavira and Adnan Darwiche The Automated Reasoning group at UCLA

IdR: an overview

Influence diagrams (ID) and Bayesian network (BN) with discrete random variables as R scripts (pure R):

- Description of the model: graph + probability distributions
- Build the graph and assign the probability distributions
- Learn Bayesian networks from data and simulate data from the model
- Inference: optimal decision policies and marginal distributions
- Mining the results for validation and sensitivity analysis

A. Einstein
"The formulation of a problem is more important than its solution"

Decision Support

Models $\rightarrow$ Evaluation $\rightarrow$ simulation $\rightarrow$ learning $\rightarrow$ queries $\rightarrow$ explantions

- Influence Diagram & Bayesian Network
  Bayes' theorem, Conditional expectation,
  Expected Utility maximization
- Node definition
  Attributes
  Arcs and
  Probabilistic dependence
  Conditional probability tables (CPT's):
  Probability distributions (matrix by rows)

R code

```
node <- function( Node=NULL, ## copy
Type=NULL, Name=NULL, Values=NULL, Preds=NULL, Pots=NULL, ## new
Mpot=NULL, Maxszpot=0, EPSILON=1e-25, trz.definition=FALSE)
```

Main functions. . .

- node: type, values, links, probabilistic dependences
- influence.diagram (chance, decision and utility variables)
- bayesian.network (chance variables)

Inference functions. . .

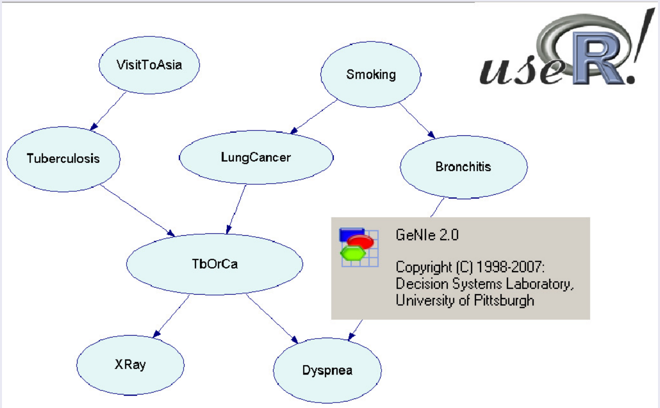
- remarc.eval and remnod.eval – ID exact evaluation
- marnod.eval and sample.eval – BN exact and approximate inference
- evid.inst – queries

Core functions (graph and probability management)...

- check.rr: DAG conditions and other properties (c.c, mady)
- bayes.i, bayes.j – implementation of Bayes rule
- conditional.expectation – combine utilities and uncertainty
- max.utility – define the optimal decision policy on every scenario

Bayesian network definition and evaluation: Asia

A "hello world" bayesian network



Bayesian network **definition** and evaluation: Asia

```
## Bayesian Network, 31-07-10

asia = list(

VisitToAsia = node( Type = "CHANCE", Name = "VisitToAsia", Values =
c("NoVisit","Visit"), Preds = c(),
Pots = matrix( data = c(
0.99, 0.01),
nrow=1,ncol=2,byrow=TRUE,dimnames=NULL)),

Smoking = node( Type = "CHANCE", Name = "Smoking", Values =
c("NonSmoker","Smoker"), Preds = c(),
Pots=matrix( data = c(
0.50, 0.50),
nrow=1,ncol=2,byrow=TRUE,dimnames=NULL)),
```

Bayesian network **definition** and evaluation: Asia

```
Tuberculosis = node( Type = "CHANCE", Name = "Tuberculosis",
Values=c("ABSENT","PRESENT"), Preds=c("VisitToAsia"),
Pots=matrix( data = c(
0.99, 0.01,
0.95, 0.05),
nrow = 2, ncol = 2, byrow = TRUE, dimnames = NULL)),

LungCancer = node( Type = "CHANCE", Name = "LungCancer",
Values=c("ABSENT","PRESENT"), Preds=c("Smoking"),
Pots=matrix( data = c(
0.99, 0.01,
0.90, 0.10),
nrow = 2, ncol = 2, byrow = TRUE, dimnames = NULL)),
```

Bayesian network **definition** and evaluation: Asia

```
Bronchitis = node( Type = "CHANCE", Name = "Bronchitis",
Values=c("ABSENT","PRESENT"), Preds=c("Smoking"),
Pots=matrix( data = c(
0.70, 0.30,
0.40, 0.60),
nrow = 2, ncol = 2, byrow = TRUE, dimnames = NULL)),

TbOrCa = node( Type = "CHANCE", Name = "TbOrCa",
Values=c("None","Sick"), Preds=c("Tuberculosis","LungCancer"),
Pots=matrix( data = c(
1.0, 0.0,
0.0, 1.0,
0.0, 1.0,
0.0, 1.0),
nrow = 4, ncol = 2, byrow = TRUE, dimnames = NULL)),
```

Bayesian network **definition** and evaluation: Asia

```
XRay = node( Type = "CHANCE", Name = "XRay",
Values=c("NORMAL","ABNORMAL"), Preds=c("TbOrCa"),
Pots=matrix( data = c(
0.95, 0.05,
0.02, 0.98),
nrow = 2, ncol = 2, byrow = TRUE, dimnames = NULL)),

Dyspnea = node( Type = "CHANCE", Name = "Dyspnea",
Values=c("ABSENT","PRESENT"), Preds=c("TbOrCa","Bronchitis"),
Pots=matrix( data = c(
0.90, 0.10,
0.20, 0.80,
0.30, 0.70,
0.10, 0.90),
nrow = 4, ncol = 2, byrow = TRUE, dimnames = NULL))
)
cat( "Bayesian network – asia: ", names(asia),"n")
```

Bayesian network **definition** and evaluation: Asia

BN represents the joint probability distribution over all variables using the chain rule and the independences expressed by the graph

Node definition

Bronchitis = node(Type = "CHANCE", Name = "Bronchitis",
Values=c("ABSENT","PRESENT"), Preds=c("Smoking"),
Pots=matrix(data = c(0.70, 0.30, 0.40, 0.60),
nrow = 2, ncol = 2, byrow = TRUE, dimnames = NULL)),

Reference

This is an example of graphical model useful in demonstrating basic concepts of Bayesian networks in diagnosis. Lauritzen, Steffen L. & Spiegelhalter, David J. (1988). Local computations with probabilities on graphical structures and their application to expert systems, Journal of the Royal Statistical Society B, 50(2):157-224.

Bayesian network definition and **evaluation**: Asia

> summary.network(asia, verbose=TRUE)

Network summary: netname, Nodes: 8

Arcs: 8, Max preds: 2, Max succs: 2

Node	grade	complexity	Predecessor
VisitToAsia	0	2	
Smoking	0	2	
Tuberculosis	1	4	VisitToAsia
LungCancer	1	4	Smoking
Bronchitis	1	4	Smoking
TbOrCa	2	8	Tuberculosis LungCancer
XRay	1	4	TbOrCa
Dyspnea	2	8	TbOrCa Bronchitis

Graphical Models for Decision Making

IdR, bn

Bayesian network definition and **evaluation**: Asia

> summary.network(asia, verbose=TRUE)

Network summary: netname, Nodes: 8

Arcs: 8, Max preds: 2, Max succs: 2

		1	2	3	4	5	6	7	8
1	VisitToAsia....	0	0	1	0	0	0	0	0
2	Smoking........	0	0	0	1	1	0	0	0
3	Tuberculosis...	0	0	0	0	0	1	0	0
4	LungCancer.....	0	0	0	0	0	1	0	0
5	Bronchitis.....	0	0	0	0	0	0	0	1
6	TbOrCa.........	0	0	0	0	0	0	1	1
7 6XRay..........	0	0	0	0	0	0	0	0	
8	Dyspnea........	0	0	0	0	0	0	0	0

Maxszpot: 18

TOTAL COMPLEX: 36 MAX COMPLEX: 8

Bayesian network definition and **evaluation**: Asia

> summary.network(marnod.eval(asia), verbose=TRUE)

Network summary: netname, Nodes: 8

Arcs: 18, Max preds: 5, Max succs: 5

Node	grade	complexity	Predecessor
VisitToAsia	1	4	Tuberculosis
Smoking	2	8	LungCancer Bronchitis
Tuberculosis	4	32	Bronchitis XRay Dyspnea TbOrCa
LungCancer	5	64	Tuberculosis Bronchitis XRay Dyspnea TbOrCa
Bronchitis	3	16	XRay Dyspnea TbOrCa
TbOrCa	0	2	
XRay	2	8	Dyspnea TbOrCa
Dyspnea	1	4	TbOrCa

Bayesian network definition and **evaluation**: Asia

> summary.network(asia, verbose=TRUE)

Network summary: netname, Nodes: 8

Arcs: 8, Max preds: 2, Max succs: 2

		1	2	3	4	5	6	7	8
1	VisitToAsia....	0	0	0	0	0	0	0	0
2	Smoking........	0	0	0	0	0	0	0	0
3	Tuberculosis...	1	0	0	1	0	0	0	0
4	LungCancer.....	0	1	0	0	0	0	0	0
5	Bronchitis.....	0	1	1	1	0	0	0	0
6	TbOrCa.........	0	0	1	1	1	0	1	1
7	XRay..........	0	0	1	1	1	0	0	0
8	Dyspnea........	0	0	1	1	1	0	1	0

Maxszpot: 114

TOTAL COMPLEX: 138 MAX COMPLEX: 64

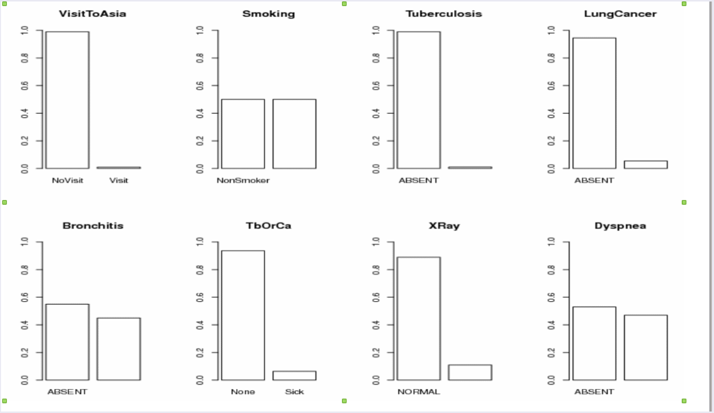
Bayesian network definition and **evaluation**: Asia

Approximate sampling evaluation
> B ¡- samnod.eval(asia)
Network summary: netname, Nodes: 8
Arcs: 8, Max preds: 2, Max succs: 2, Maxszpot: 18
TOTAL COMPLEX: 36, MAX COMPLEX: 8
SAMPLE EVALUATION: netname, size: 36

- mpd: 0.9897 0.0103 Node: VisitToAsia
- mpd: 0.4994 0.5006 Node: Smoking
- mpd: 0.9897 0.0103 Node: Tuberculosis
- mpd: 0.9478 0.0522 Node: LungCancer
- mpd: 0.5444 0.4556 Node: Bronchitis
- mpd: 0.9381 0.0619 Node: TbOrCa
- mpd: 0.1050 0.8950 Node: XRay
- mpd: 0.6050 0.3905 Node: Dyspnea

Bayesian network definition and **evaluation**: Asia

Marginal distribution



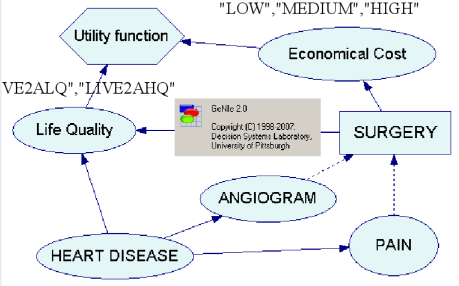
Bayesian network definition and evaluation: Asia

Exact marginalization

- 0.99 0.01 ::–n–:: VisitToAsia ——- NoVisit Visit ::–p–:: Tuberculosis
- 0.50 0.50 ::–n–:: Smoking ——- NonSmoker Smoker ::–p–:: LungCancer Bronchitis
- 0.989 0.0104 ::–n–:: Tuberculosis ——- ABSENT PRESENT ::–p–:: Bronchitis XRay Dyspnea TbOrCa
- 0.945 0.0550 ::–n–:: LungCancer ——- ABSENT PRESENT ::–p–:: Tuberculosis Bronchitis XRay Dyspnea TbOrCa
- 0.550 0.450 ::–n–:: Bronchitis ——- ABSENT PRESENT ::–p–:: XRay Dyspnea TbOrCa
- 0.946 0.054 ::–n–:: TbOrCa ——- None Sick ::–p–::
- 0.901 0.099 ::–n–:: XRay ——- NORMAL ABNORMAL ::–p–:: Dyspnea TbOrCa
- 0.528 0.471 ::–n–:: Dyspnea ——- ABSENT PRESENT ::–p–:: TbOrCa

Influence diagram definition and evaluation: ByPass

ID represents a decision process under uncertainty with a decision sequence and preferences (utility) over the results



Node types: decision, chance and utiltiy; Arc types: informative and conditional; Regular ID: oriented, no cycles, defined decision sequence; Normalized ID: normalized potentials and normalized utility function

Influence diagram **definition** and evaluation: ByPass

```
## Influence Diagram 03-09-08

bypass = list(

PAIN = node( Type = "CHANCE", Name = "PAIN", Values =
c("ABSENT","PRESENT"), Preds = c("HEARTDISEASE"),
Pots = matrix( data = c(0.80, 0.20, 0.70, 0.30),
nrow = 2, ncol = 2, byrow = TRUE, dimnames = NULL)),

ANGIOGRAM = node( Type = "CHANCE", Name = "ANGIOGRAM", Values =
c("NEGATIVE","POSITIVE"), Preds = c("HEARTDISEASE"),
Pots = matrix( data = c(0.95, 0.05, 0.15, 0.85),
nrow = 2, ncol = 2, byrow = TRUE, dimnames = NULL)),

HEARTSURGERY = node( Type = "DECISION", Name = "HEARTSURGERY",
Values = c("NO","YES"), Preds=c("PAIN","ANGIOGRAM"),
Pots = matrix( data = c(1.0), dimnames = list("phase","HEARTSURGERY"))),
```

Influence diagram **definition** and evaluation: ByPass

HEARTDISEASE = node(Type="CHANCE", Name = "HEARTDISEASE", Values=c("ABSENT","PRESENT"), Preds = c(),
Pots = matrix(data = c(0.85, 0.15),
nrow = 1,ncol = 2,byrow = TRUE,dimnames = NULL)),

LIFEQ = node(Type="CHANCE", Name = "LIFEQ", Values = c("DEAD","LIVE2ALQ","LIVE2AHQ"), Preds = c("HEARTDISEASE","HEARTSURGERY"),
Pots = matrix(data = c(0.02, 0.08, 0.90, 0.09, 0.29, 0.62, 0.15, 0.30, 0.55, 0.17, 0.23, 0.60),
nrow = 4, ncol = 3, byrow = TRUE, dimnames = NULL)),

ECONOMICALC = node(Type="CHANCE", Name = "ECONOMICALC", Values = c("LOW","MEDIUM","HIGH"), Preds = c("HEARTSURGERY"),
Pots = matrix(data = c(0.70, 0.25, 0.05, 0.05, 0.35, 0.60),
nrow=2,ncol=3,byrow=TRUE,dimnames=NULL)),

Influence diagram **definition** and evaluation: ByPass

UTILITY = node(Type="UTILITY", Name = "UTILITY", Values = c(0.0,1.0),
Preds =c("LIFEQ","ECONOMICALC"),
Pots = matrix(data = c(1.0, 0.90, 0.70, 0.80, 0.50, 0.10, 1.40, 1.50, 1.80),
nrow = 9,ncol = 1,byrow = TRUE,dimnames = list(NULL, c("UTILITY"))))
)

cat("Influence Diagram – bypass: ", names(bypass),"n")

Influence diagram **definition** and evaluation: ByPass

Code for decision and utility nodes

- HEARTSURGERY = node(Type = "DECISION", Name = "HEARTSURGERY", Values=c("NO","YES"), Preds=c("PAIN","ANGIOGRAM"), Pots=matrix(data = c(1.0), dimnames = list("phase","SURGERY"))),
- UTILITY = node(Type="UTILITY", Name="UTILITY", Values=c(0.0,1.0), Preds=c("LIFEQ","ECONOMICALC"), Pots=matrix(data=c(1.0, 0.90, 0.70, 0.80, 0.50, 0.10, 1.40, 1.50, 1.80), nrow=9,ncol=1,byrow=TRUE, dimnames=list(NULL, c("UTILITY"))))

Influence diagram definition and **evaluation**: ByPass

Evaluation output is an optimal decision table for every decision

;Decision:	HEARTSURGERY
;Preds utility node:	UTILITY $<$ PAIN ANGIOGRAM HEARTSURGERY $>$
File:	dec-HEARTSURGERY ;
S:	10 HEARTSURGERY 2 ;
Val:	NO YES ;
Att:	200 PAIN 2 ;
Val:	ABSENT PRESENT ;
Att:	300 ANGIOGRAM 2 ;
Val:	NEGATIVE POSITIVE ;
Att:	400 SURGERY 2 ;
Val:	NO YES ;

Influence diagram definition and **evaluation**: ByPass

Evaluation output is an optimal decision table for every decision
Max utility in bold font

PAIN	ANGIOGRAM	SURGERY	Utility
ABSENT	NEGATIVE	→NO	**0.74673**
ABSENT	NEGATIVE	YES	0.64070
ABSENT	POSITIVE	→NO	**0.65233**
ABSENT	POSITIVE	YES	0.64598
PRESENT	NEGATIVE	→NO	**0.74453**
PRESENT	NEGATIVE	YES	0.64083
PRESENT	POSITIVE	NO	0.63965
PRESENT	POSITIVE	→YES	**0.64668**

Influence diagram definition and **evaluation**: ByPass

Evaluation output is an optimal decision table for every decision

Optimal policy (Rules):

Pain Absent & Angiogram Negative then Surgery No
Pain Absent & Angiogram Positive then Surgery No
Pain Present & Angiogram Negative then Surgery No
Pain Present & Angiogram Positive then Surgery Yes

Influence diagram definition and evaluation: ByPass; **explanation**

KBM2L: Kowledge Base Matrix to List
Fernandez del Pozo, J. A., C. Bielza, and M. Gómez, "A List-Based Compact Representation for Large Decision Tables Management", European Journal of Operational Research, vol. 160, no. 3, pp. 638-662, 2005.

PAIN	ANGIOGRAM	SURGERY	Utility
ABSENT	NEGATIVE	NO	0.746733
ABSENT	NEGATIVE	YES	0.640708
ABSENT	POSITIVE	NO	0.652331
ABSENT	POSITIVE	YES	0.645981
PRESENT	NEGATIVE	NO	0.744533
PRESENT	NEGATIVE	YES	0.640831
PRESENT	POSITIVE	NO	0.639655
PRESENT	POSITIVE	YES	0.646689

Table - Multidimensional Matrix

Influence diagram definition and evaluation: ByPass; **explanation**

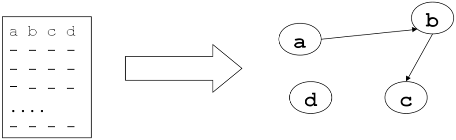
KBM2L: Kowledge Base Matrix to List
Bielza, C., J. A. Fernandez del Pozo, and P. Lucas, "Explaining Clinical Decisions by Extracting Regularity Patterns", Decision Support Systems

- KBM2L:
  $< (Present, Negative), No| < (Present, Positive), Yes|$
- Explanation:
  - Surgery No $< -$ (Pain Absent) OR (Angiogram Negative)
  - Sugery Yes $< -$ (Pain Present) AND (Angiogram Positive)
- The best explanation is available using the most concise list; How? Searching the proper permutation of the attibutes (and domains) on the table! Also useful for conditional probability tables.

Learning Bayesian Networks

- Learning a marginal / naive / generic Bayesian network model for Classification
- Structure learning and probability model estimation

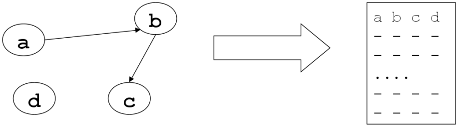
lbn function code:
function (filename, tab0, nettype = "mbayes", noise = NULL, class.name = NULL, maxnode.input = -1, laplace.correction = FALSE, time.out = 300, trz.probability = FALSE)



Simulating Bayesian Networks

- Simulating a data set from the model for Inference
- Sampling the model $P(abcd) = P(a)P(d)P(b|a)P(c|b)$

sbn function code:
function (bn.file, tab0, ssz = N, nettype = "xbayes", class.name = NULL, trz.probability = FALSE)



Future lines of research

- More general decision networks (continuous variables, several utility nodes, non sequential decision nodes,. . .)
- Alternatives to the (large) conditional probability tables (linear models) and utility tables (multiattribute utility functions)
- Implementation of an R package for KBM2L (java) → KBMR
- Evaluation and learning algorithms from data
- Complex queries, MPE, MAP
  Fernandez del Pozo, J. A., and C. Bielza, "Dealing with Complex Queries in Decision Support Systems", Data & Knowledge Engineering, vol. 70, pp. 167-181, 2011
- We are interesting on paralell evaluation of huge models, using packages like snow, i.e. very large decision sequences

¿Remarks and Questions?

IDA 2015

12/01/2015, CIG