Inteligencia Artificial
Razonamiento probabilístico

Primavera 2007

profesor: Luigi Ceccaroni

Redes Bayesianas

- A simple, graphical notation for conditional independence assertions and hence for compact specification of full joint distributions.

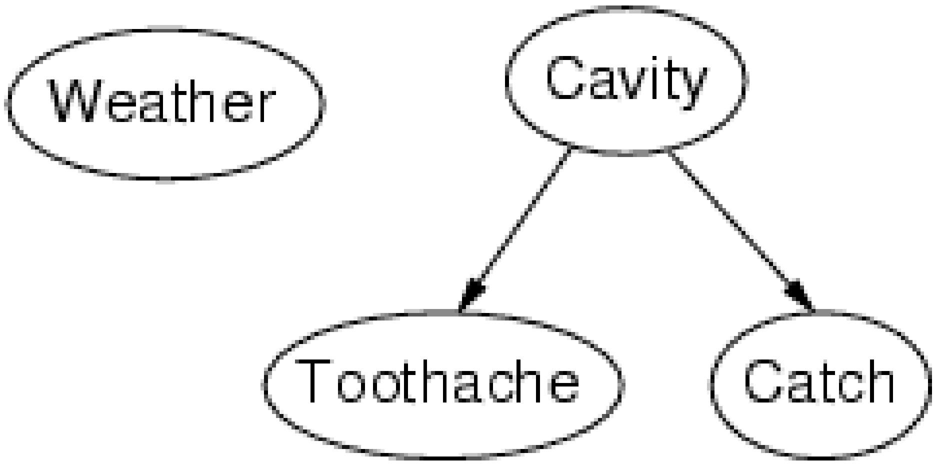
- Syntax:
  - a set of nodes, one per variable
  - a directed, acyclic graph (links ≈ "directly influences")
  - a conditional distribution for each node given its parents:

$$\mathbf{P}(X_i \mid Parents(X_i))$$

- In the simplest case, conditional distribution represented as a conditional probability tables (CPTs) giving the distribution over X_i for each combination of parent values.

Example

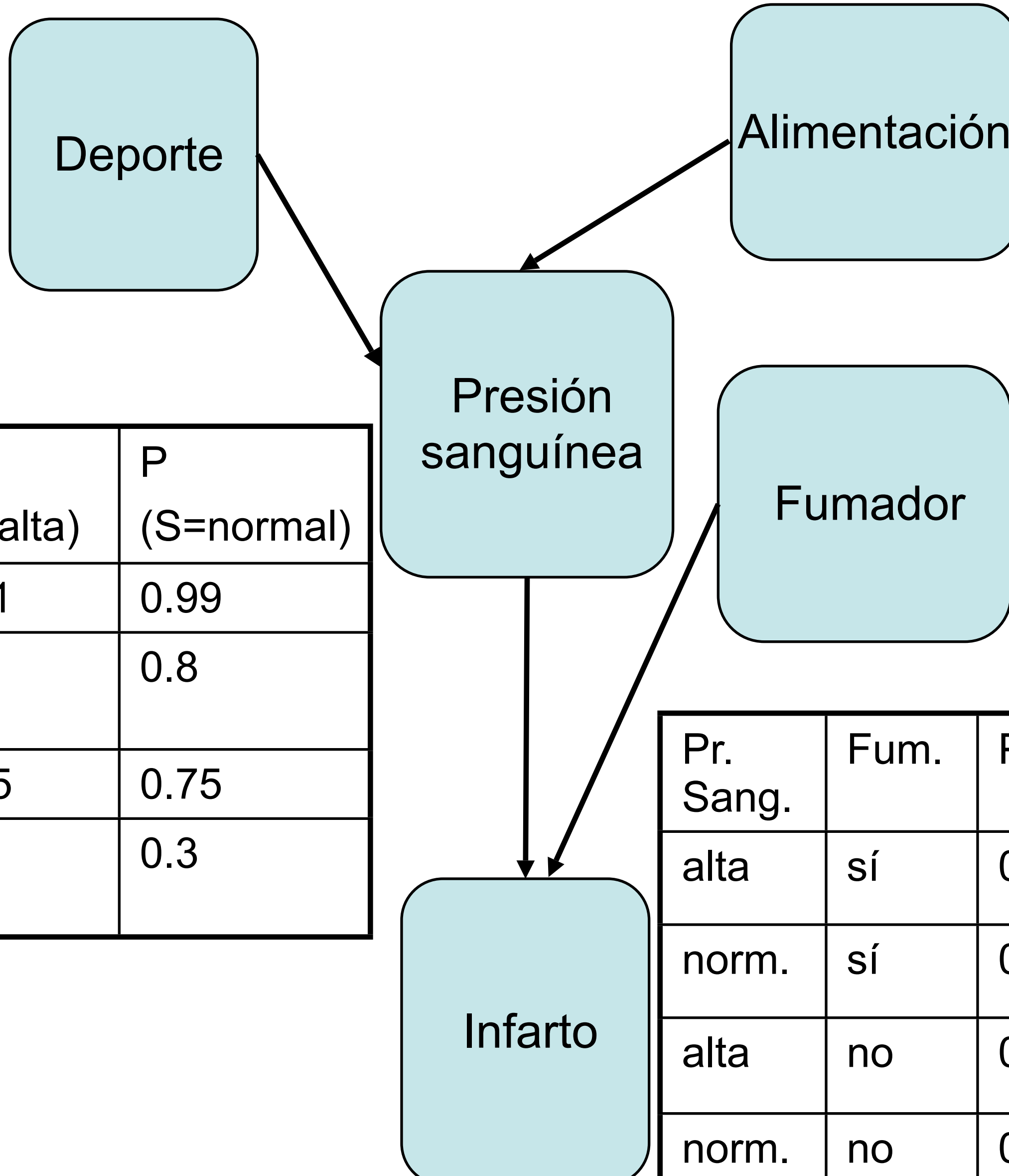
- Topology of network encodes conditional independence assertions:



- *Weather* is independent of the other variables.
- *Toothache* and *Catch* are conditionally independent given *Cavity.*

Ejemplo

Deporte	P(D)
sí	0.1
no	0.9



Alimentación	P(A)
equilibrada	0.4
no equilibrada	0.6

Alim.	Deporte	P (S=alta)	P (S=normal)
eq.	sí	0.01	0.99
no eq.	sí	0.2	0.8
eq.	no	0.25	0.75
no eq.	no	0.7	0.3

Fumador	P(F)
sí	0.4
no	0.6

Pr. Sang.	Fum.	P(I=sí)	P(I=no)
alta	sí	0.8	0.2
norm.	sí	0.6	0.4
alta	no	0.7	0.3
norm.	no	0.3	0.7

Compactness

- A CPT for Boolean X_i with k Boolean parents has 2^k rows for the combinations of parent values.

- Each row requires one number p for X_i = *true* (the number for X_i = *false* is just $1-p$).

- If each variable has no more than k parents, the complete network requires $O(n \cdot 2^k)$ numbers.

- I.e., grows linearly with n, vs. $O(2^n)$ for the full joint distribution.

Semantics

The full joint distribution is defined as the product of the local conditional distributions:

$$\mathbf{P}(X_1, \ldots, X_n) = \pi_{i=1}^{n} \mathbf{P}(X_i \mid Parents(X_i))$$

e.g., $\mathbf{P}(sp \wedge d{=}balanced \wedge p{=}high \wedge \neg sm \wedge \neg h)$

$= \mathbf{P}(sp)\, \mathbf{P}(d{=}balanced)\, \mathbf{P}(p{=}high \mid sp, d{=}balanced)\, \mathbf{P}(\neg sm)\, \mathbf{P}(\neg h \mid p{=}high, \neg sm)$

Constructing Bayesian networks

- 1. Choose an ordering of variables $X_1, \dots, X_n$
- 2. For i = 1 to n
  - add X_i to the network
  - select parents from $X_1, \dots, X_{i-1}$ such that

$$\mathbf{P}(X_i \mid Parents(X_i)) = \mathbf{P}(X_i \mid X_1, \dots X_{i-1})$$

This choice of parents guarantees:

$$\mathbf{P}(X_1, \dots, X_n) = \pi_{i=1}^{n} \mathbf{P}(X_i \mid X_1, \dots, X_{i-1}) \text{ (chain rule)}$$

$$= \pi_{i=1}^{n} \mathbf{P}(X_i \mid Parents(X_i)) \text{ (by construction)}$$

Summary

- Bayesian networks provide a natural representation for (causally induced) conditional independence.
- Topology + CPTs = compact representation of joint distribution.
- Generally easy for domain experts to construct.