

RAZONAMIENTO APROXIMADO

EN LA INTELIGENCIA ARTIFICIAL

Introducción a la Inteligencia Artificial
LCC

Ana Casali

REALIDAD

El **conocimiento** que necesitamos para desarrollar un Sistema basado en Conocimiento tiene muchas veces las siguientes características:

**NO ES DEL TODO
CONFIABLE**

IMPRECISO

INCOMPLETO

CONTRADICTORIO

Causas de inexactitud



Generalmente no es del todo confiable
(falta de evidencias, excepciones)

Suele ser incompleta a la hora de tomar decisiones
(faltan datos provenientes de mediciones, análisis)

Diferentes fuentes pueden ser conflictivas,
redundantes, subsumidas

El lenguaje usado para transmitirla es
inherentemente impreciso, vago

REALIDAD

Las personas con esas fuentes de conocimiento, dotadas de esas características, razonamos y muchas veces concluimos ...



**CAPACIDAD DE RAZONAR
APROXIMADAMENTE**

PROBLEMA

Como modelizamos estas características del *conocimiento*, de modo de poder:



- **REPRESENTARLO**
- **UTILIZARLO**

REALIDAD

La lógica clásica es un buen modelo para formalizar cualquier razonamiento basado en información certera (V o F)



NECESITAMOS OTROS FORMALISMOS

REALIDAD

El desarrollo de la IA ha incentivado el estudio de formalismos que son alternativos o complementarios a la lógica clásica



**INVESTIGACION Y DESARROLLO DE
OTROS FORMALISMOS**

Ejemplos

Como representar en una BC ...

Si el paciente tiene el Signo1 y el Signo2 entonces el diagnóstico en el 75% de los casos es D1 y en el 40% de los casos es D2

Y si se tiene...

Un paciente que evidencia Signo1 en un 80% y Signo2 en un 55%

QUE SE PUEDE INFERRIR ???

Ejemplos

Como representar en una BC ...

Si la humedad es alta, la presión es baja y está muy nublado, entonces lloverá.

Y si se tiene...

Que la humedad es del 75%, la presión es 1002hp y esta nublado.

QUE SE PUEDE INFERRIR ???

INGENIERIA DEL CONOCIMIENTO



PROBLEMA

Tomar decisiones y realizar procesos de razonamiento cuando el conocimiento del dominio involucrado tiene distintas características, puede ser:

INCIERTO

IMPRECISO

INCOMPLETO

NO-MONOTONO

CONOCIMIENTO INCIERTO

El conocimiento se expresa mediante predicados precisos pero **no podemos establecer el valor de verdad** de la expresión

Ejemplos:

- Es posible que A pese más de 10kg
A pesa más de 10kg CF
- Creo que el auto era rojo
El auto es rojo CF

CONOCIMIENTO INCERTO

Cuando no podemos establecer la verdad o falsedad de la información



Debemos evaluar la : PROBABILIDAD
POSIBILIDAD
NECESIDAD/PLAUSIBILIDAD
GRADO DE CERTEZA...

De que la información sea verdadera

MEDIDA DE INCERTIDUMBRE $(\text{EVENTO}) = \text{VALOR} / \text{VALORES}$



bivaluado

CONOCIMIENTO IMPRECISO

El conocimiento cuenta con predicados o cuantificadores vagos (no precisos)

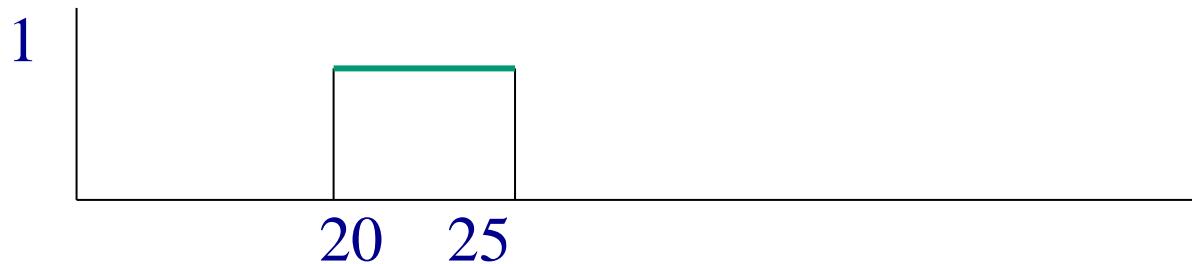
Ejemplos:

- Pedro tiene entre 20 y 25 años.
- Juan es joven
- Mucha gente juega al fútbol
- El espectáculo es para gente grande.

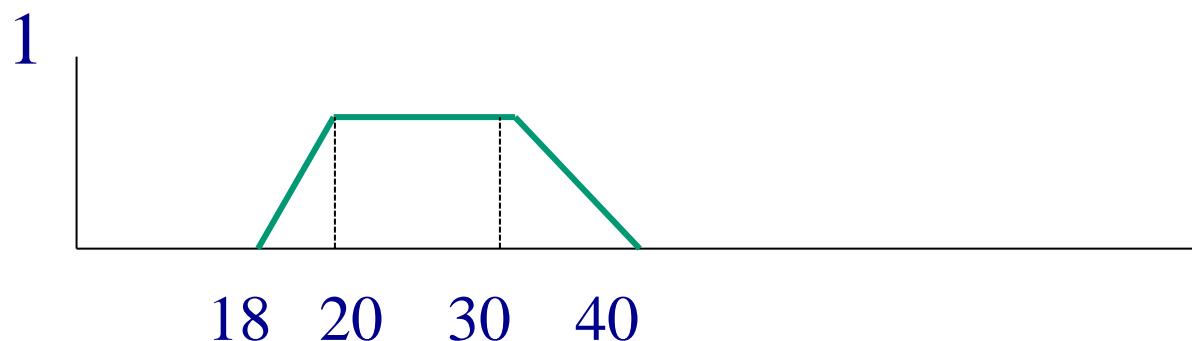
CONOCIMIENTO IMPRECISO

Ejemplos:

- Pedro tiene entre 20 y 25 años.



- Juan es joven



CONOCIMIENTO IMPRECISO

Si la variable X toma valores en S

- Proposiciones precisas

$$\{p: "X \text{ es } s" / s \in S\}$$

- Proposiciones imprecisas

$$\{p: "X \text{ es } r" / r \subset S\}$$

- * Imprecisa - no borrosa

Si r es un conjunto clásico

- * Imprecisa - borrosa (fuzzy)

Si r es un conjunto borroso (fuzzy)

CONOCIMIENTO INCOMPLETO

Se debe tomar decisiones a partir de
información incompleta o parcial.

Esto se suele manejar a través de supuestos
o valores por defecto.

➤ Ejemplo:

Si el paciente tiene S1, S2 y S3 entonces
tiene una infección a Bacteria

S3 ???

CONOCIMIENTO NO-MONOTONO

La información recibida a partir de distintas fuentes o en diferentes momentos *es conflictiva y cambiante.*

➤ Ejemplo:

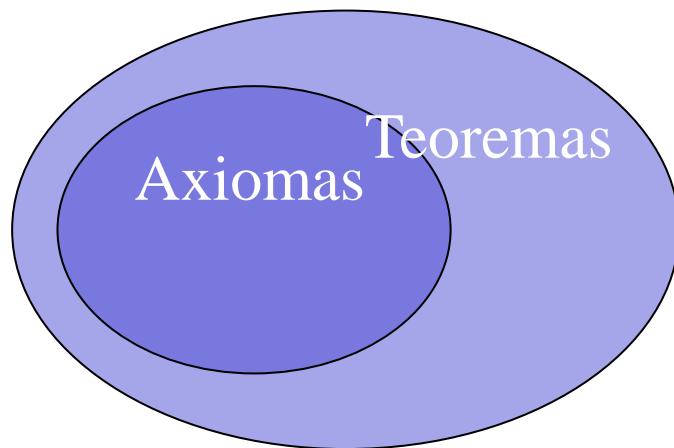
Si el vuelo nº 1340 sale en forma puntual y no tiene escalas técnicas arribará a Madrid a las 8 hs

1º Supongo no-escala técnica y concluyo arribará a Madrid a las 8 hs

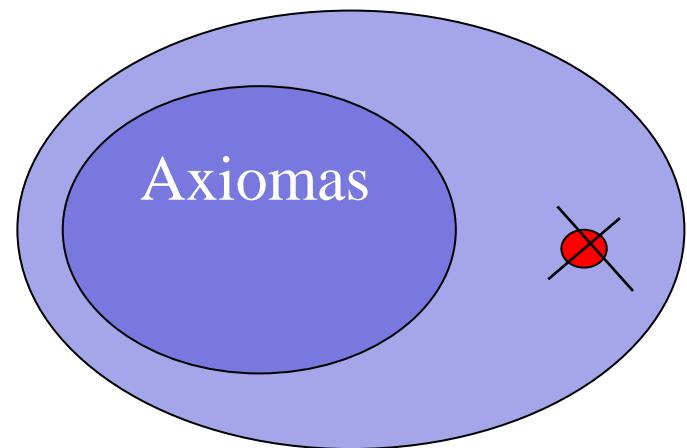
2º Aviso de escala técnica, debo revisar la conclusión del horario de arribo.

LÓGICAS NO-MONÓTONAS

Lógicas clásicas: Monótonas



Lógicas No-Monótonas



RAZONAMIENTO APROXIMADO (RA)

Trata como

- REPRESENTAR
- COMBINAR y
- REALIZAR INFERENCIAS

con conocimiento impreciso y/o incierto

RA: Esquema general en sistemas basados en reglas de producción

Hipótesis :

- Si X es A entonces Y es B (λ)
- X es A*

Conclusión

- Y es B* ???

REGLAS IMPRECISAS: A y/ o B son imprecisos

REGLA INCIERTA: λ Grado de certeza

REGLAS HIBRIDAS: Problema complejo

RA: Distintos modelos

- MODELOS PROBABILISTICOS
- MODELO EVIDENCIAL
- MODELO POSIBILISTICO

- Todos tratan la incertidumbre en un sistema de reglas
- Sólo el modelo posibilístico puede tratar la imprecisión → FUZZY LOGIC.

MODELOS PROBABILISTICOS



Probabilidad - Axiomas

- $P: \text{PROP} \rightarrow [0,1]$
- $P(V) = 1$ y $P(F) = 0$
- $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$
- Propiedad $P(\neg A) = 1 - P(A)$

Probabilidad - Conceptos

- Probabilidad a priori o incondicional
 - $P(A)$ o $P(X=S)$
- Variables aleatorias: X, Y
- Dominio: $\{x_1, x_2, \dots, x_n\}$ exhaustivo y excluyente
- Probabilidad condicional:
 - $P(A/B)$ $P(X/Y)$ tabla valores $P(X=x_i/Y=y_k)$
 - $P(A/B) = P(A \wedge B) / P(B)$

Distribución de Probabilidad Conjunta

	DolorD	\neg DolorD
Caries	0.04	0.06
\neg Caries	0.01	0.89

- $P(\text{Caries} \vee \text{DolorD}) = 0.04 + 0.06 + 0.01 = 0.11$
- $P(\text{Caries} / \text{DolorD}) =$

$$\begin{aligned}&= P(\text{Caries} \wedge \text{DolorD}) / P(\text{DolorD}) = \\&= 0.04 / 0.04 + 0.01 = 0.8\end{aligned}$$

➤ Problema exponencial con la cantidad de variables

La regla de Bayes

- $P(B/A) = P(A/B) * P(B) / P(A)$



Es la base de muchos de los sistemas
de inferencia probabilística
 $(A \rightarrow B) \dots P(B/A)$

RA: Modelos probabilísticos

- ✓ Modelo utilizado en Prospector
(Duda-Hart' 81)
- ✓ Modelo utilizado en Mycin
(Shortliffe-Buchanan' 75-84)
- ✓ Redes Bayesianas (Redes de Creencias - Pearl'86)

MYCIN

EL MODELO DE RAZONAMIENTO APROXIMADO PARA MANEJO DE LA INCERTIDUMBRE, BASADO EN LOS CFs

- ✓ Si bien tiene poco fundamento teórico
 - Alguna base en teoría de probabilidades
 - Regla de combinación de Dempster-Shafer
- ✓ Ha sido muy utilizado en el desarrollo de SE e implementado en algunos Shells

REDES BAYESIANAS

RA: Redes Bayesianas

- Para representar la dependencia que existe entre determinadas variables, en aplicaciones complejas, se utiliza una estructura de datos conocida como

*Red Bayesiana, Red de creencias,
Red Probabilística o Red causal.*

- Esta estructura sirve para especificar de manera concisa la distribución de probabilidad conjunta.

RA: Redes Bayesianas

- REDES DE RELACIONES PROBABILISTICAS ENTRE PROPOSICIONES (variables aleatorias) RELACIONADAS SEMANTICAMENTE (relaciones causales) REPRESENTADAS MEDIANTE UN GDA



REDES BAYESIANAS

NODOS



PROPOSICIONES (variable o

conjunto de variables)

ARCOS



RELACIONES CAUSALES

(X ejerce influencia directa sobre Y)

PESO DE ARCOS



PROBABILIDAD CONDICIONAL

(Tabla de Probabilidad Condicional)

RA: Redes Bayesianas

➤ Hay que establecer:

✓ Topología de la red



A los expertos les resulta
fácil determinar las dependencias entre conceptos

✓ Probabilidades condicionales



Tarea más compleja

(datos estadísticos, subjetivos, utilizar otras técnicas
aprendizaje automatizado)

RA: Redes Bayesianas

➤ Topología de la red:

Podría considerarse como una base de conocimientos abstractos, válida en una gran cantidad de escenarios diversos



Representa la estructura general de los procesos causales del dominio

RA: Redes Bayesianas

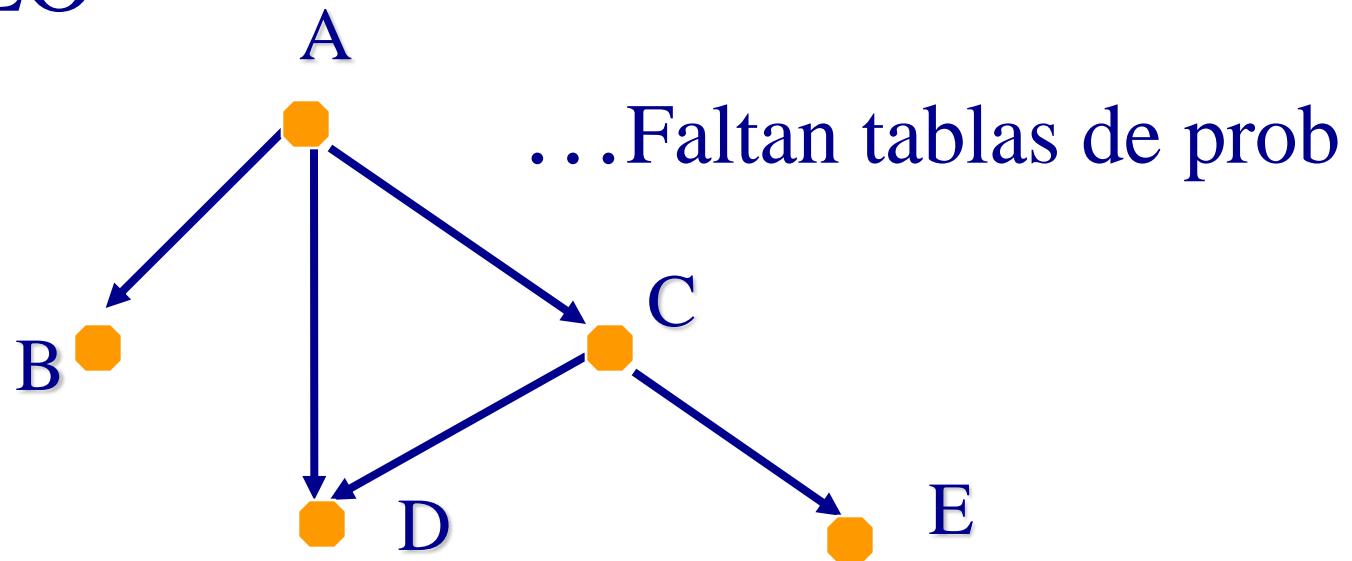
- La **incertidumbre** inherente a los distintos enlaces (relaciones causales) representan las situaciones no representadas explícitamente.



Las **probabilidades** resumen un conjunto de posibles circunstancias en que pueden ser verdaderas (falsas) las variables de un nodo.

RA: Redes Bayesianas

✓ EJEMPLO



Del grafo, que representa las relaciones causales, se puede sacar la distribución conjunta:

$$p(A, B, C, D, E) = P(E / C) P(D / A,C) P(C / A) P(B / A) P(A)$$

RA: Redes Bayesianas

- En general, es posible calcular cada una de las entradas de la distribución conjunta desde la información de la red

$$P(x_1, \dots, x_n) = \prod_{i=1,n} P(x_i / \text{Padres}(x_i))$$

RA: Redes Bayesianas

➤ EJEMPLO (Norvig & Russell / Judea Pearl)

Una casa tiene una alarma que se activa ante intento de robo, pero puede activarse ante temblores (el escenario es en Los Angeles).

Dos vecinos, Juan y María se han ofrecido a llamar al dueño de la casa al trabajo, si escuchan la alarma. Juan a veces confunde el sonido de la alarma con otros sonidos, pero llama de todos modos y María a veces no la escucha por otras fuentes de sonido que tiene encendida (TV, Música).

RA: Redes Bayesianas

➤ EJEMPLO

Objetivo: Realizar distintas de inferencias

Con la evidencia de quien ha llamado
y quien no



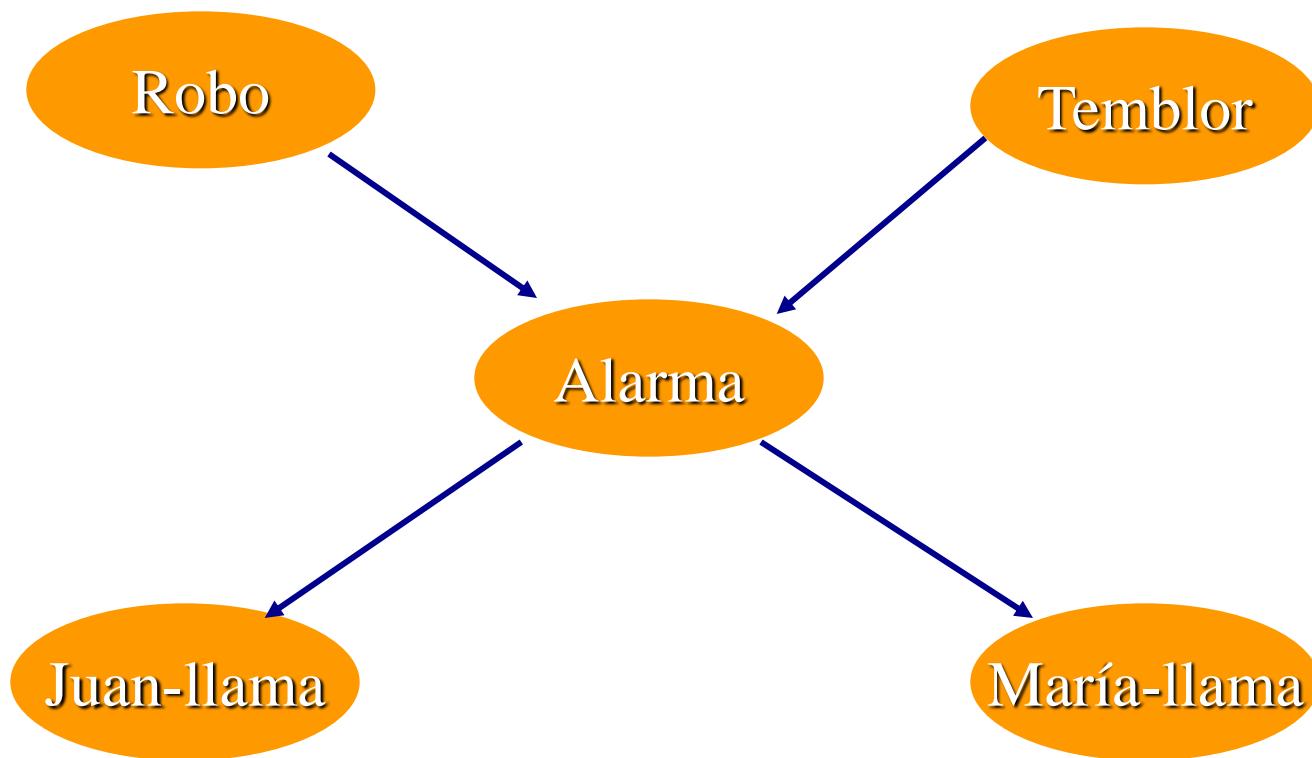
Cual es la Probabilidad de robo????

$$P(R/\neg J, M)$$

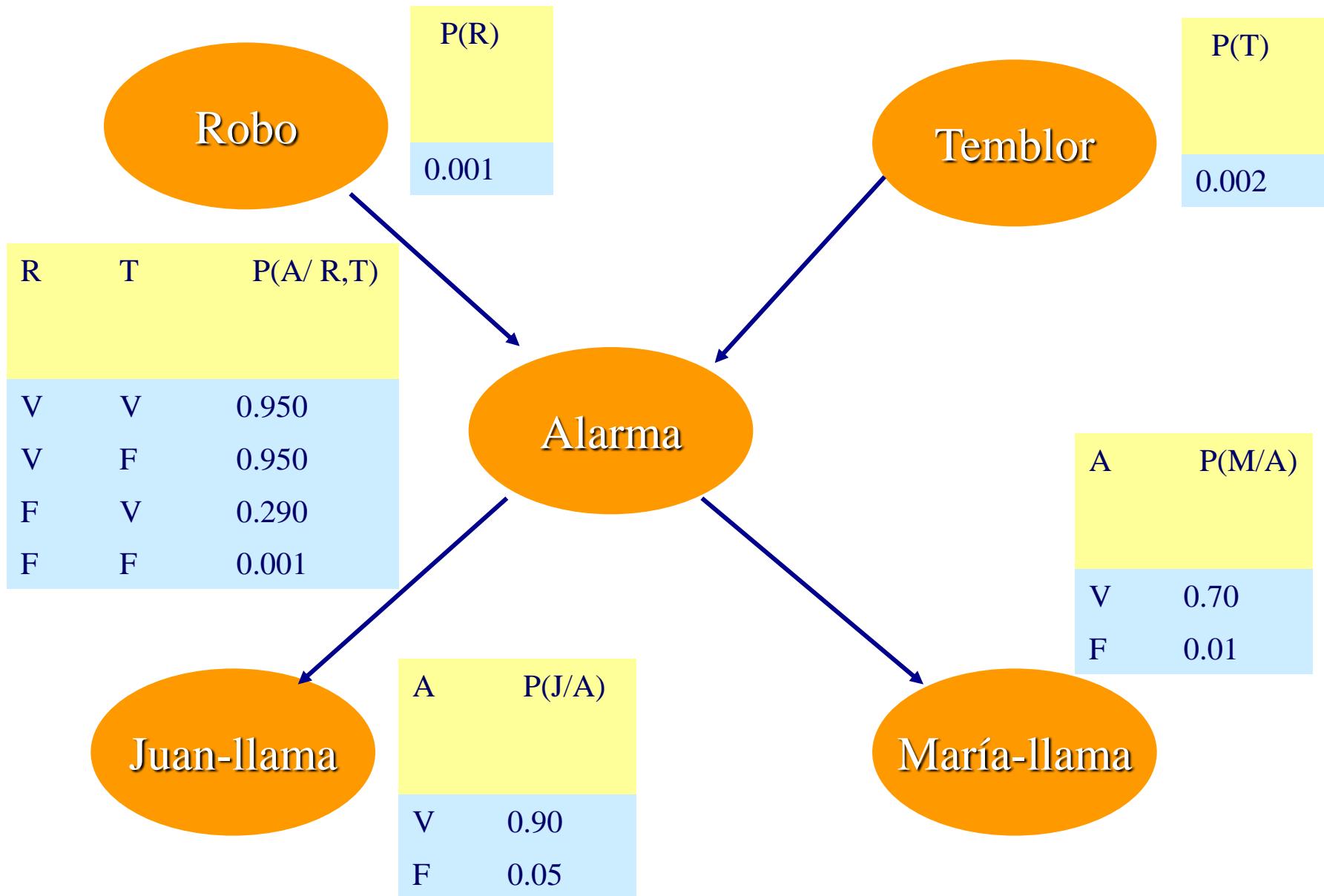
RA: Redes Bayesianas

➤ EJEMPLO

TOPOLOGIA DE LA RED



EJEMPLO



RA: Redes Bayesianas

➤ EJEMPLO

Como ejemplo podemos calcular la probabilidad del evento de que suene la alarma, sin que se haya producido robo ni temblor, habiendo llamado Juan solamente:

$$P(J \wedge \neg M \wedge A \wedge \neg R \wedge \neg T) = P(J/A) P(\neg M/A) \\ P(A/\neg R \wedge \neg T) P(\neg R) P(\neg T) = 0.00062$$

- Si la Red Bayesiana es una representación de la probabilidad conjunta, sirve para responder consultas del dominio
- $$P(R / J) ??? = 0.16$$

RA: Redes Bayesianas

- EJEMPLO
- Si la Red Bayesiana es una representación de la probabilidad conjunta, sirve para responder consultas del dominio $P(R / J) ??? = 0.16$

$$P(R / J) = P(R \wedge J) / P(J)$$

Cómo uso la red para computar estas probabilidades llevándolas a las probabilidades conjuntas???

RA: Redes Bayesianas

- EJEMPLO
- Si la Red Bayesiana es una representación de la probabilidad conjunta, sirve para responder consultas del dominio $P(R / J) ??? = 0.16$

$$P(R / J) = P(R \wedge J) / P(J)$$

$$P(R \wedge J) = \sum_{x \in \{M, -M\}, y \in \{A, -A\}, z \in \{T, -T\}} P(R, J, x, y, z)$$

$$P(R, J, A, M, T) = P(J/A) P(M/A) P(A/R, T) P(R) P(T)$$

así cada probabilidad conjunta...

RA: Redes Bayesianas

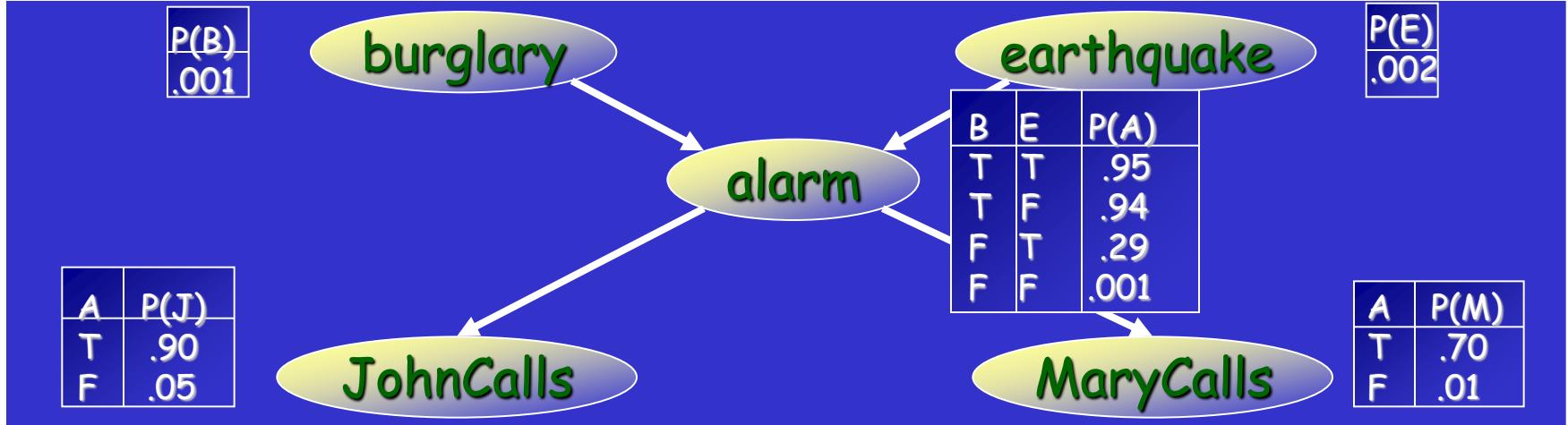
- EJEMPLO
- Si la Red Bayesiana es una representación de la probabilidad conjunta, sirve para responder consultas del dominio $P(R / J) = 0.16$

Luego $P(J) = P(R \wedge J) + P(\neg R \wedge J)$

$$P(\neg R \wedge J) = \sum_{x \in \{M, \neg M\}, y \in \{A, \neg A\}, z \in \{T, \neg T\}} P(\neg R, J, x, y, z)$$

Se procede de forma igual al cómputo de $P(R \wedge J)$

Belief Networks as Rule Systems



Burglary (0.01) ←
Earthquake (0.02) ←

Alarm (0.95) ← Burglary ∧ Earthquake
 Alarm (0.94) ← Burglary ∧ \neg Earthquake
 Alarm (0.29) ← \neg Burglary ∧ Earthquake
 Alarm (0.001) ← \neg Burglary ∧ \neg Earthquake

JohnCalls (0.90) ← Alarm
 JohnCalls (0.05) ← \neg Alarm

MaryCalls (0.70) ← Alarm
 MaryCalls (0.01) ← \neg Alarm

Doesn't add anything ...but shows that you need many rules to represent the full Bayesian Net

In many cases you may not have all this information!

RA: Redes Bayesianas

- INDEPENDENCIA: Se hace explícita mediante la separación de grafos.
- SE CONSTRUYE INCREMENTALMENTE por el experto agregando objetos y relaciones.
- Los arcos no deben considerarse estáticos, representan restricciones sobre la certeza de los nodos que unen

$p(A / B)$ cuantifica la certeza de $B \rightarrow A$
si lo que se conoce es una evidencia e de que B es cierto  $p(A/B,e)$

RA: Redes Bayesianas

Inferencias: Belief revision

- ✓ Consiste en encontrar la asignación global que maximice cierta probabilidad
- ✓ Puede usarse para tareas explicatorias/diagnóstico
- ✓ Básicamente a partir de cierta evidencia E, nuestra tarea es encontrar un conjunto de hipótesis que constituyan la mejor explicación de las evidencias
- ✓ Encontrar asignaciones a los nodos N₁...N_j / $(P(E, N_1, \dots, N_j))$ sea máxima. Por ejemplo llamó Juan, x,y z w? / $P(J, x, y, z, w)$ sea máxima

RA: Redes Bayesianas

Inferencias: Belief updating

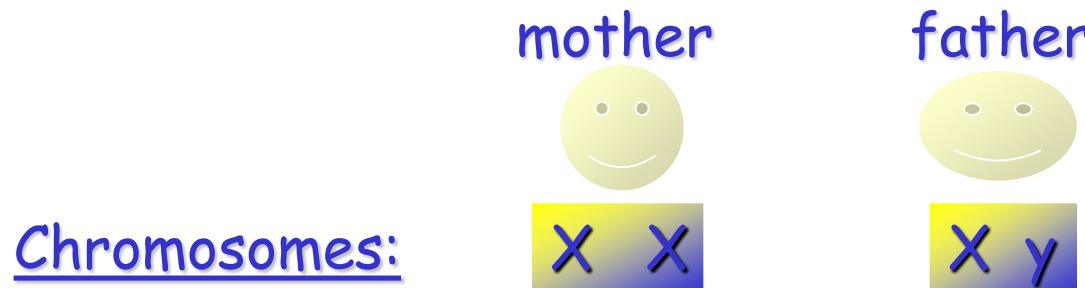
- ✓ Consiste en determinar la mejor instanciación de una variable, dada una evidencia.
- ✓ Es la actualización de probabilidades de un nodo dadas un conjunto de evidencias:
 - ✓ $(P(N_i/E_1, \dots, E_n))$
- ✓ Ejemplo: determinar la probabilidad de robo sabiendo que Juan llama y María llama.

$P(R/J, M)$  actualiza el valor de $P(R)$

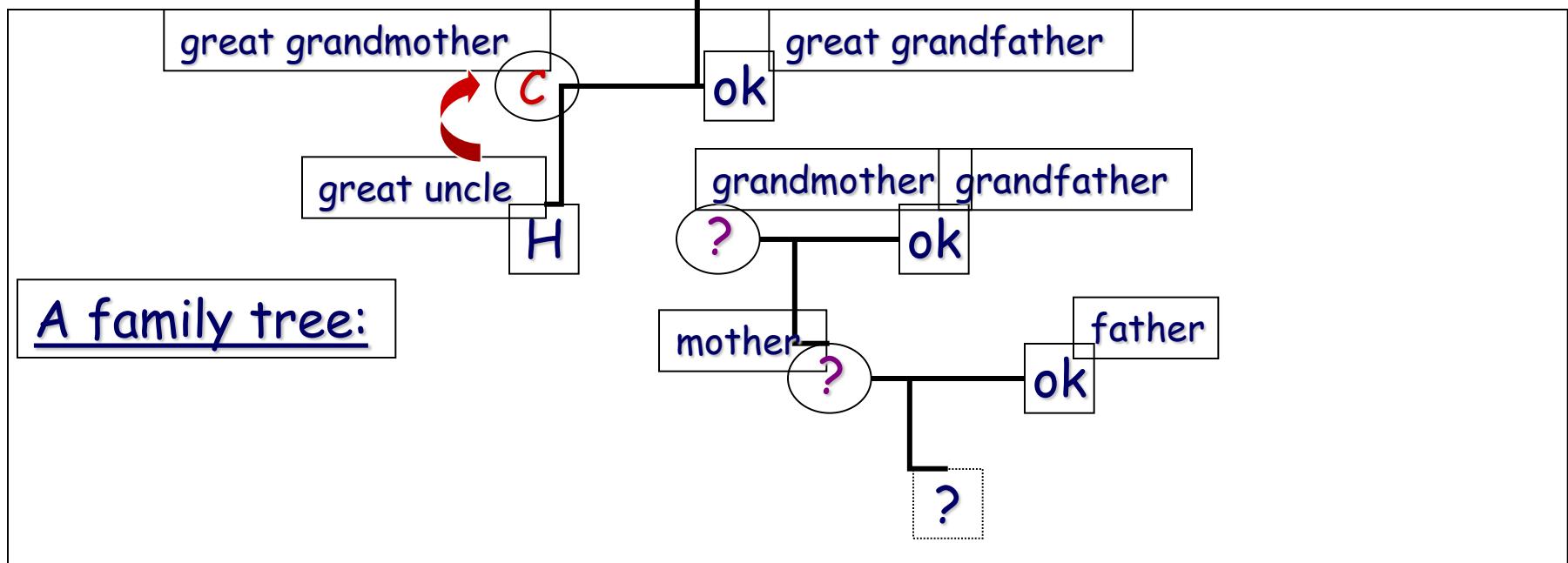
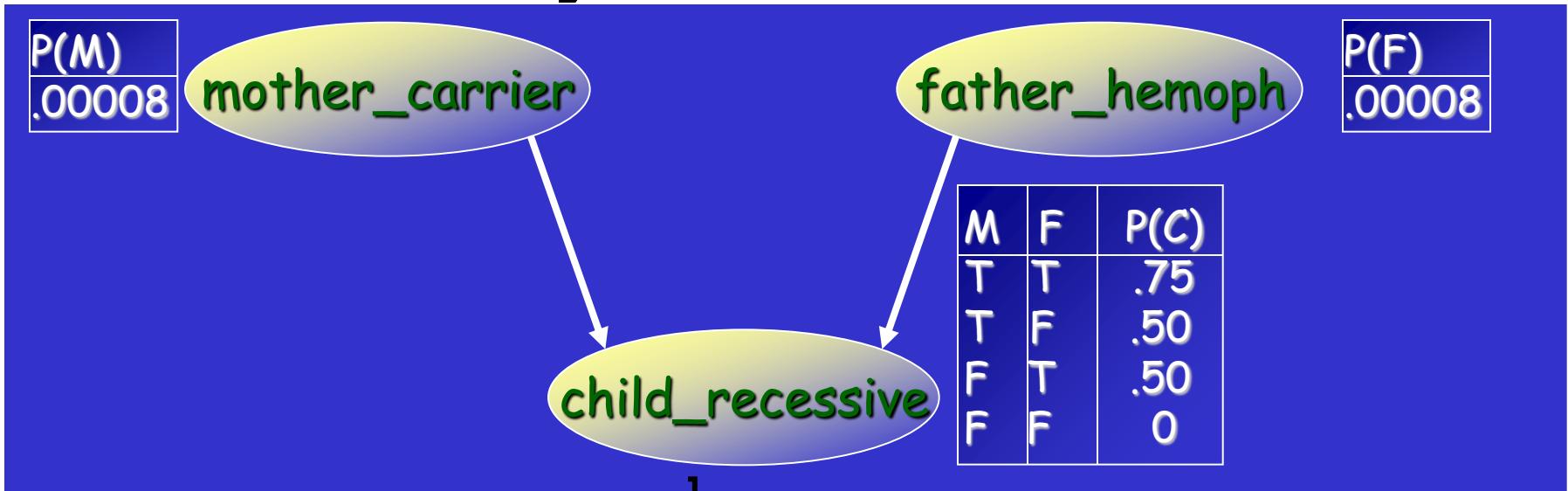
An application: GENINFER

A couple is expecting a child.

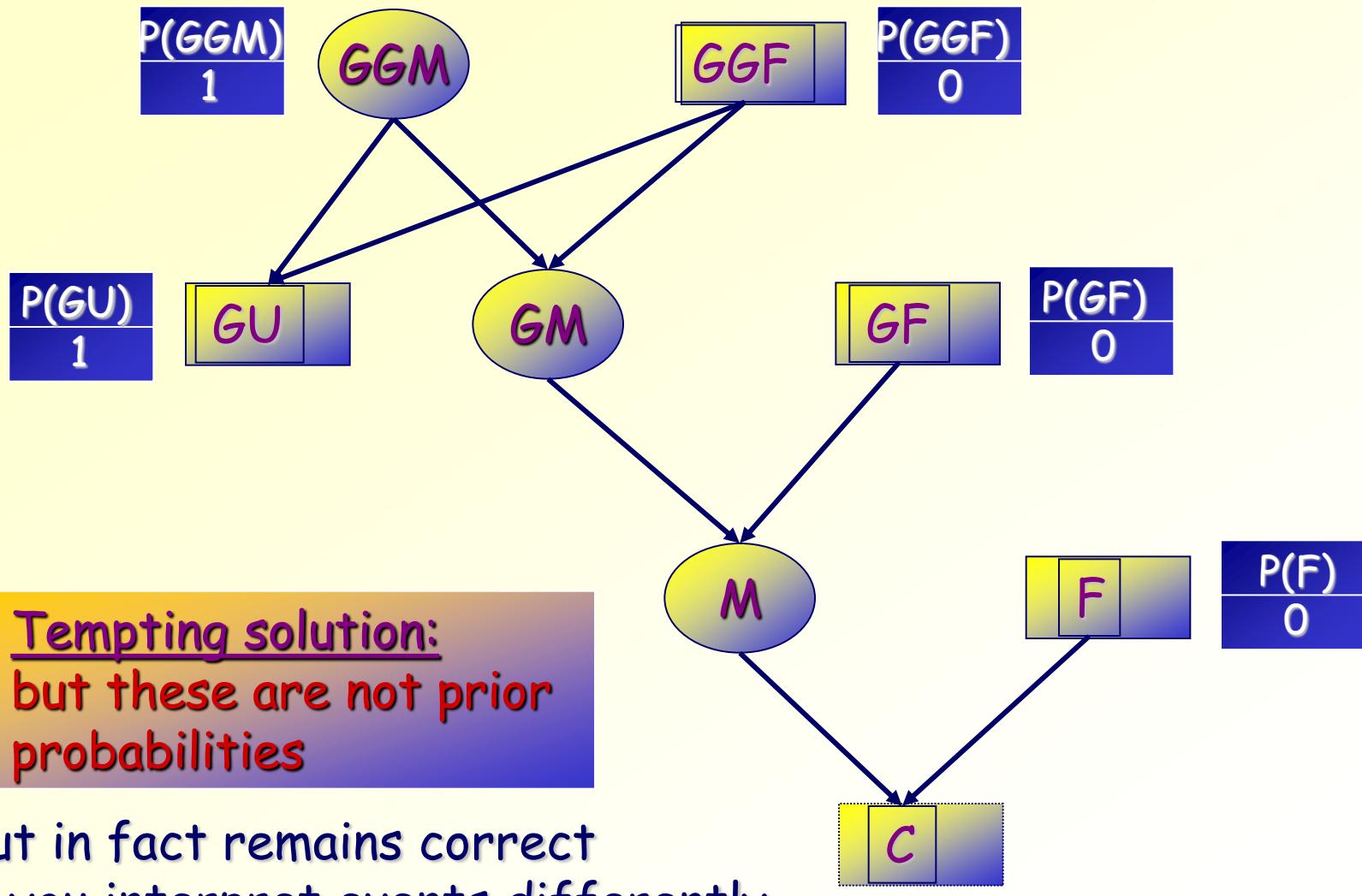
- The (expecting) mother has a hemophiliac risk
- determine the probability of hemophiliac for the child
- Hemophiliac disease is genetically determined:
 - Due to a defected X chromosome



The Bayesian Network:



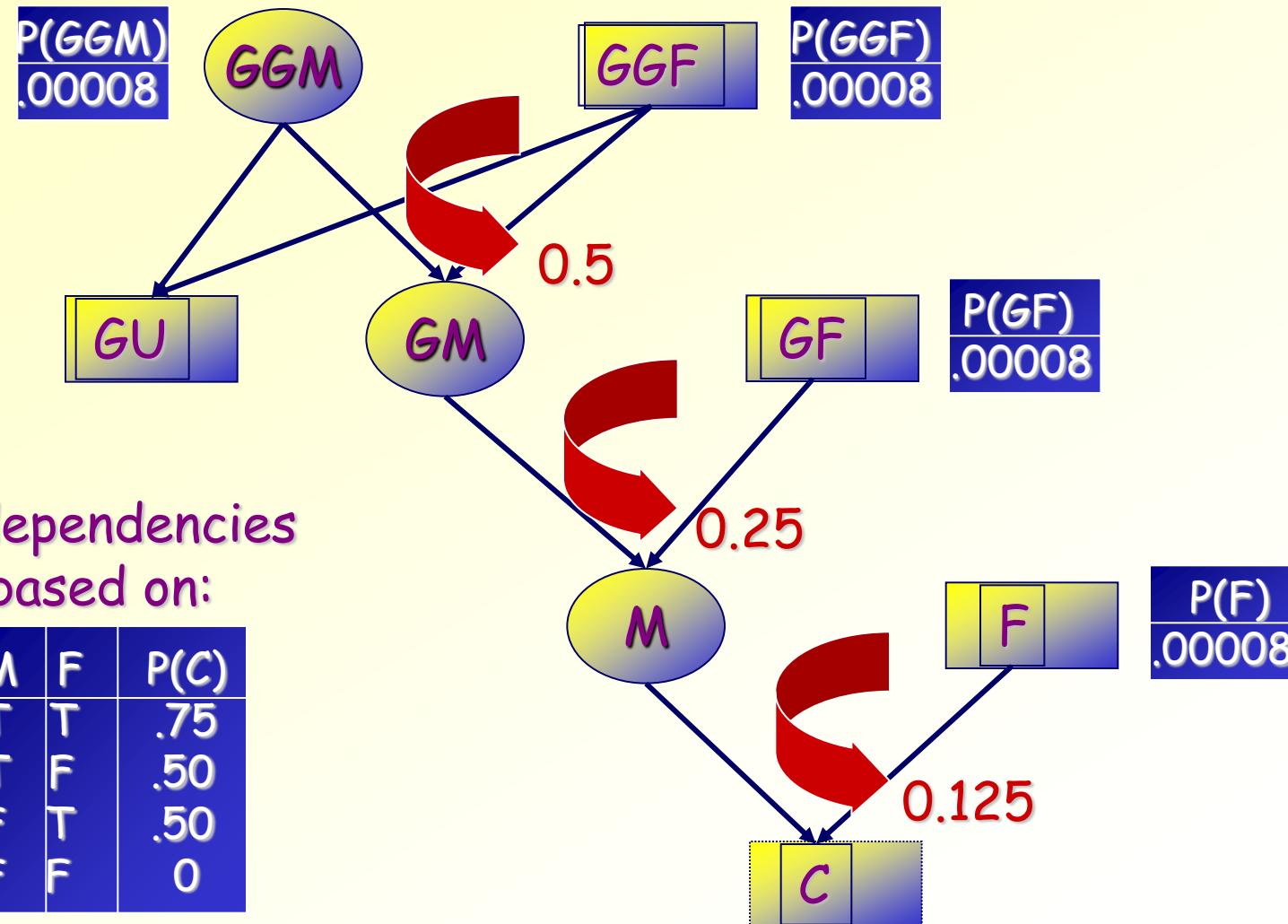
Expanding to full network:



Expanding to full network (2)

Compute: $P(GGM | GU \wedge \neg GGF) = 1$

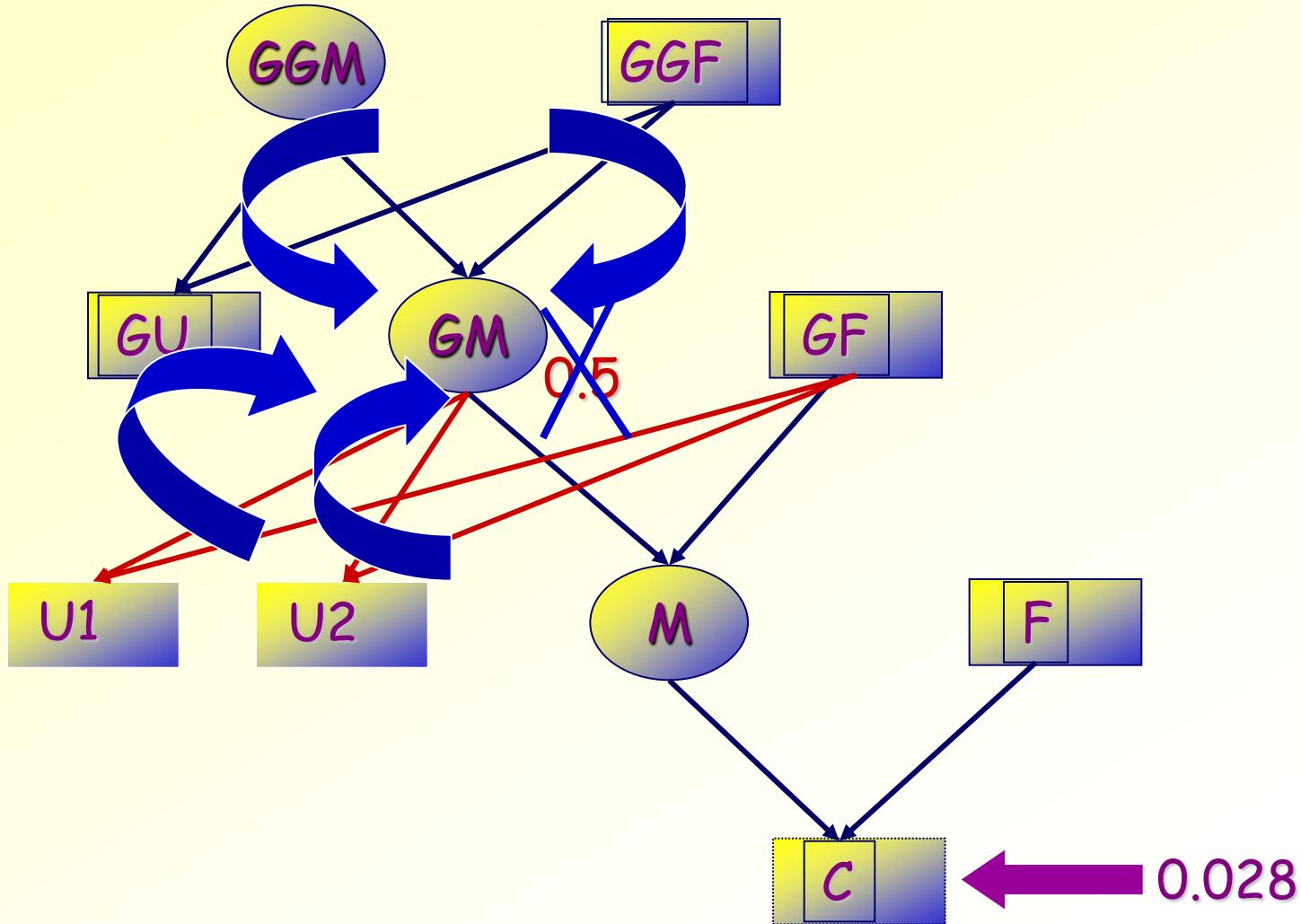
Compute: $P(GM | GGM \wedge \neg GGF) = 0.5$, etc.



And if there are uncles?

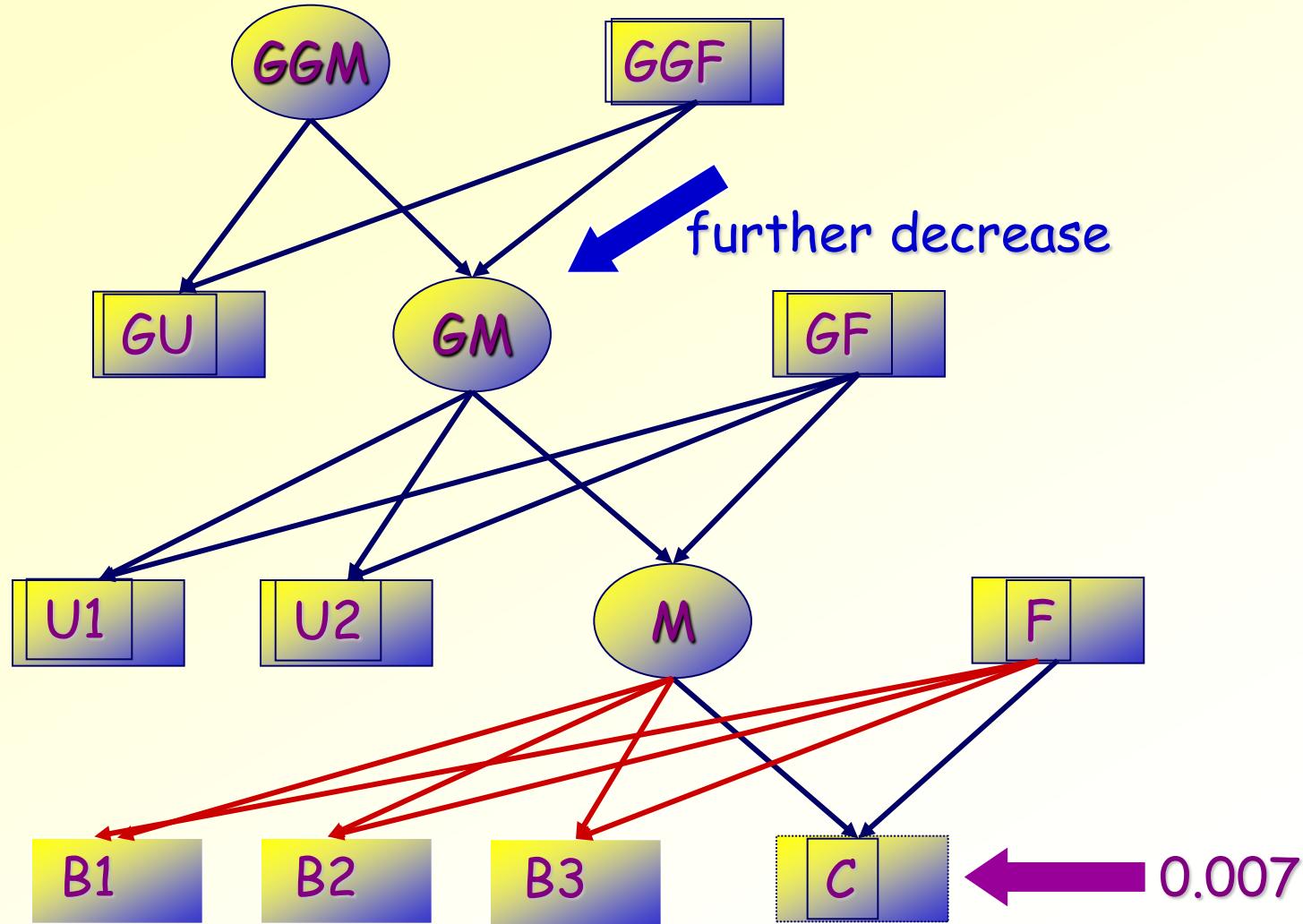
Recompute: $P(GM | GGM \wedge \neg GGF \wedge \neg U1 \wedge \neg U2)$

Propagate the information to Mother and Child



And brothers?

Probability under additional condition of 3 healthy bothers:



MODELOS PROBABILISTICOS

- Problema de las asignaciones de probabilidad (estadísticas o evaluaciones subjetivas?)
- Las Redes Bayesianas son modelos más cercanos a un modelo probabilístico puro y permite la representación explícitas de las dependencias del dominio en la red.
 - *Razonamiento con redes: JavaBayes*
 - *Aprender Redes: herramientas como R*