
Inteligencia Artificial

Redes Bayesianas

Jorge Luis Guevara Diaz

www.jorge.sistemasyservidores.com

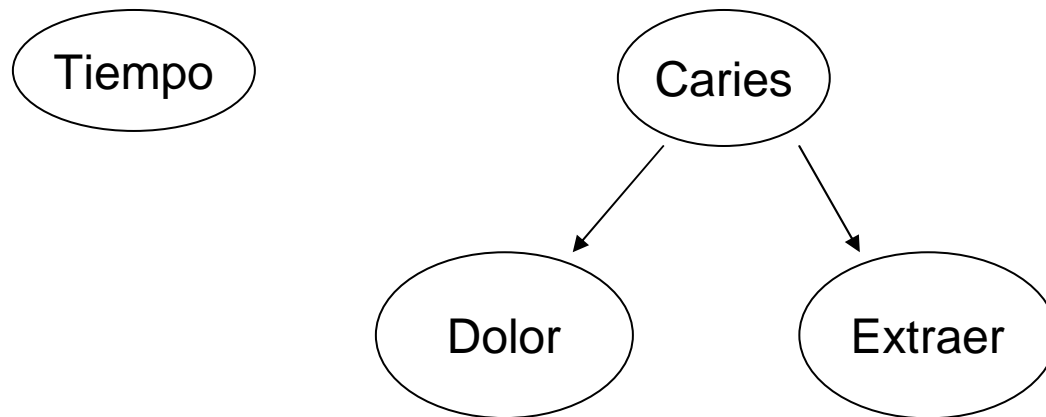
Redes Bayesianas

- Es una notación gráfica, para mostrar independencias condicionales y por lo tanto es una compacta especificación de una distribución conjunta completa
- Sintaxis
 - Un conjunto de nodos : un nodo por variable
 - Un grafo dirigido sin ciclos : (enlazado con las influencias directas)
 - Una distribución condicional para cada nodo dados sus padres $P(X_i | \text{Padres}(X_i))$

En el caso mas simple la distribución condicional es representada por una **tabla de probabilidad condicional** (CPT), dada la distribución sobre X_i para cada combinación de los valores de los padres

Redes Bayesianas : ejemplo

- La topología de la red codifica las afirmaciones de independencia condicional

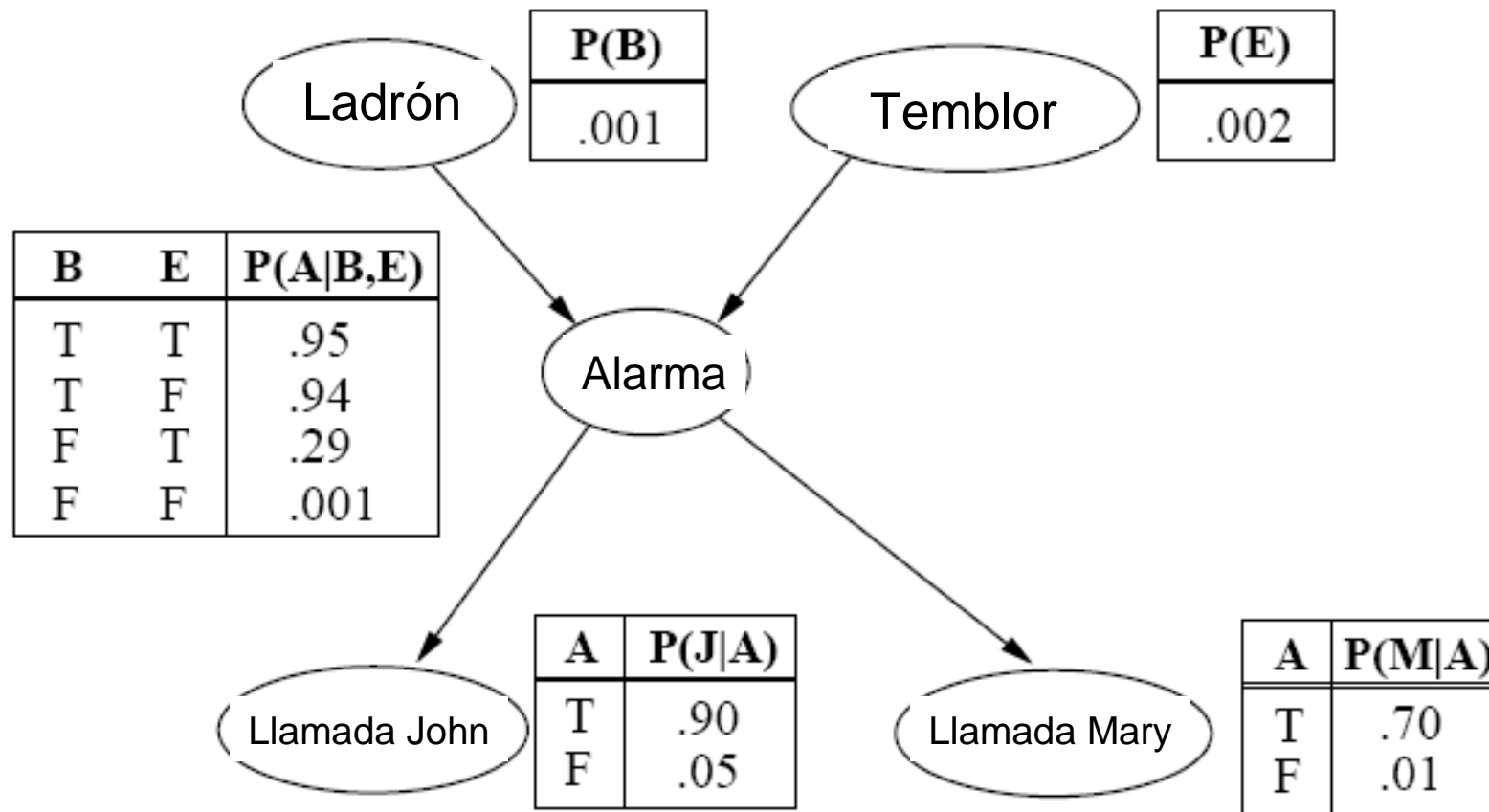


- Tiempo es independiente de otras variables
- Dolor y Extraer son condicionalmente independientes dado Caries

Redes Bayesianas : ejemplo

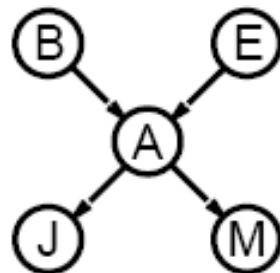
- Estoy en el trabajo, mi vecino John me llama para decirme que mi alarma esta sonando, pero mi vecina Mary no lo hace, algunas veces la alarma suena con pequeños temblores, Hay un ladrón en mi casa?
- Variables :
Ladron = B, Temblor = E, Alarma = A, Llamada de John =J,
LlamadaMary = M
- La topología de la red refleja el conocimiento causal
 - Un ladrón puede activar la alarma
 - Un temblor puede activar la alarma
 - La alarma causa que Mary llame
 - La alarma causa que John llame

Redes Bayesianas : ejemplo



Redes Bayesianas : Compactación

- Una CPT para una variable booleana X_i con k Padres booleanos tiene 2^k para la combinación de los valores de los padres
- Cada columna requiere un valor p para $X_i=true$, siendo el numero de $X_i=false$ $1-p$
- Si cada variable no tiene mas que k padres, la red completa requiere solo $O(n2^k)$ numeros
- Crece linealmente con n vs $O(2^n)$ para la distribucion conjunta completa
- Para la red del ladrón $1+1+4+2+2 = 10$ numeros vs $(2^5-1=32)$



Redes Bayesianas : Semántica Global

- La semántica global define la distribución conjunta completa como el producto de distribuciones condicionales locales

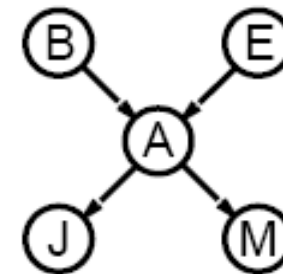
$$P(x_1, \dots, x_n) = \prod_{i=1}^n P(x_i | \text{parents}(X_i))$$

e.g., $P(j \wedge m \wedge a \wedge \neg b \wedge \neg e)$

$$= P(j|a)P(m|a)P(a|\neg b, \neg e)P(\neg b)P(\neg e)$$

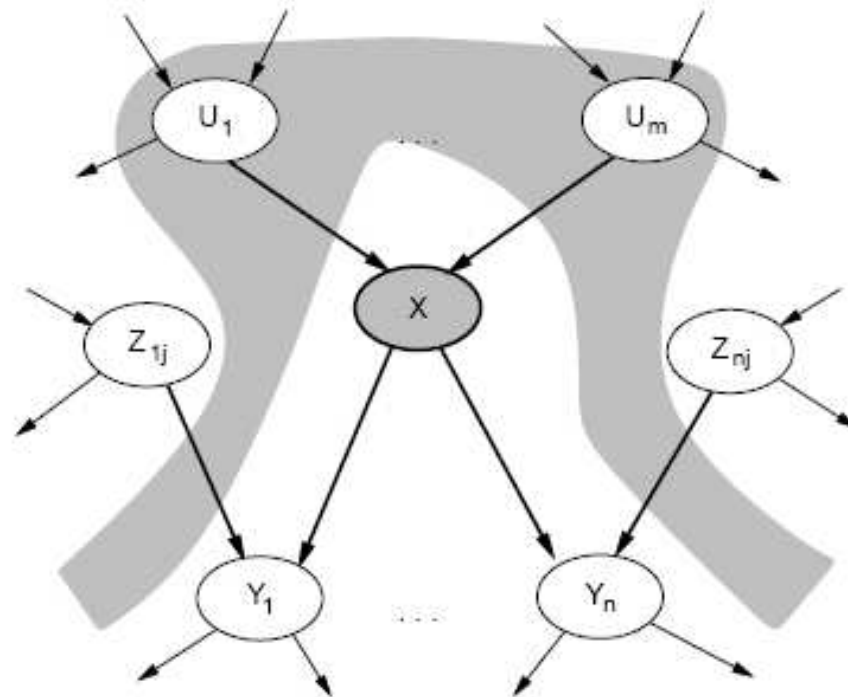
$$= 0.9 \times 0.7 \times 0.001 \times 0.999 \times 0.998$$

$$\approx 0.00063$$



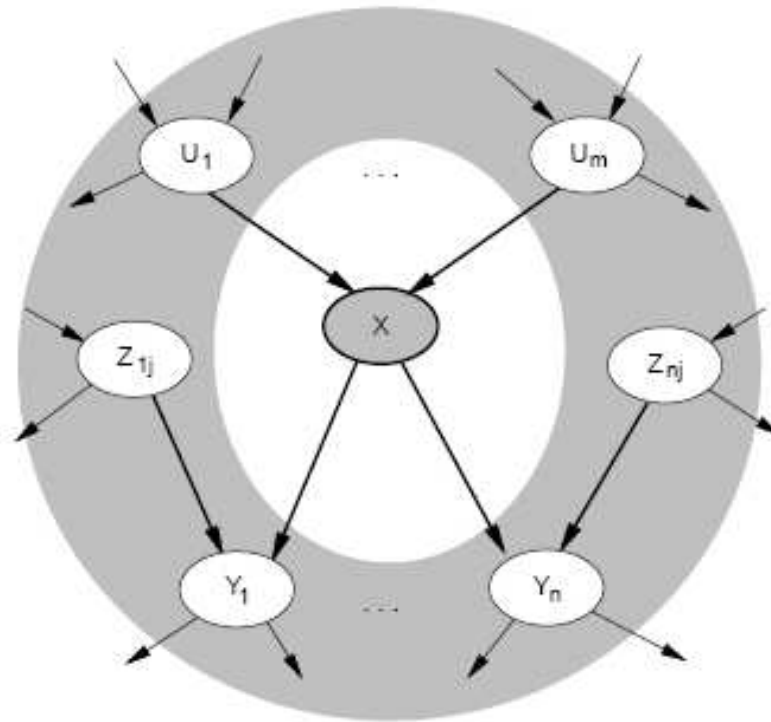
Redes Bayesianas : Semántica Local

- Cada nodo es condicionalmente de sus no descendientes dado sus padres



Redes Bayesianas : Capa de Markov

- Cada nodo es condicionalmente independiente de todos los otros dado su **Capa de Markov (Markov Blanket)**



Redes Bayesianas : ¿Cómo construirlas?

- Es necesario un método de tal manera que una serie de afirmaciones locales verificables de independencia condicional garantice la semántica global requerida

1. Escoger un orden de variables X_1, \dots, X_n

2. For $i = 1$ to n

Agregar X_i a la red

Escoger los padres de X_1, \dots, X_{i-1} de tal manera que

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

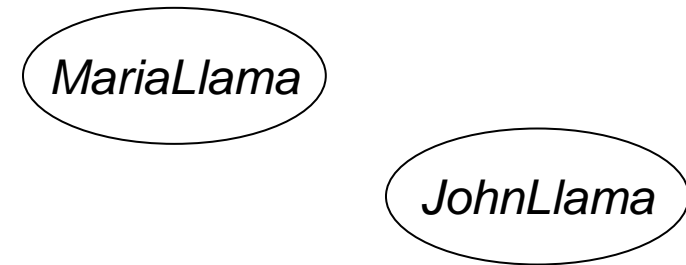
1. Esta eleccion de los padres garantiza la semántica global

$$\begin{aligned} P(X_1, \dots, X_n) &= \prod_{i=1}^n P(X_i | X_1, \dots, X_{i-1}) \\ &= \prod_{i=1}^n P(X_i | Parents(X_i)) \end{aligned}$$

Redes Bayesianas : ¿Cómo construirlas?

Ejemplo

- Supongamos el siguiente orden *M, J, A, B, E*

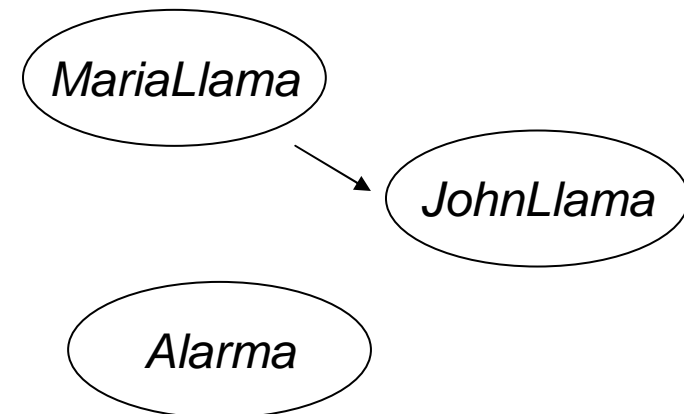


$$P(J | M) = P(J)?$$

Redes Bayesianas : ¿Cómo construirlas?

Ejemplo

- Supongamos el siguiente orden M, J, A, B, E



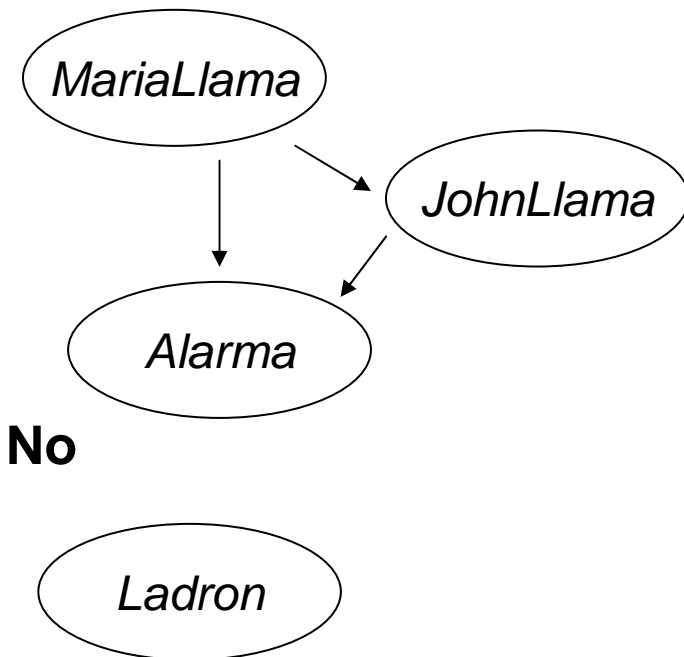
$$P(J | M) = P(J)? = \mathbf{No}$$

$$P(A | J, M) = P(A | J)? \quad P(A | J, M) = P(A)?$$

Redes Bayesianas : ¿Cómo construirlas?

Ejemplo

- Supongamos el siguiente orden *M, J, A, B, E*



$P(J | M) = P(J)$? =No

$P(A | J, M) = P(A | J)$? $P(A | J, M) = P(A)$? No

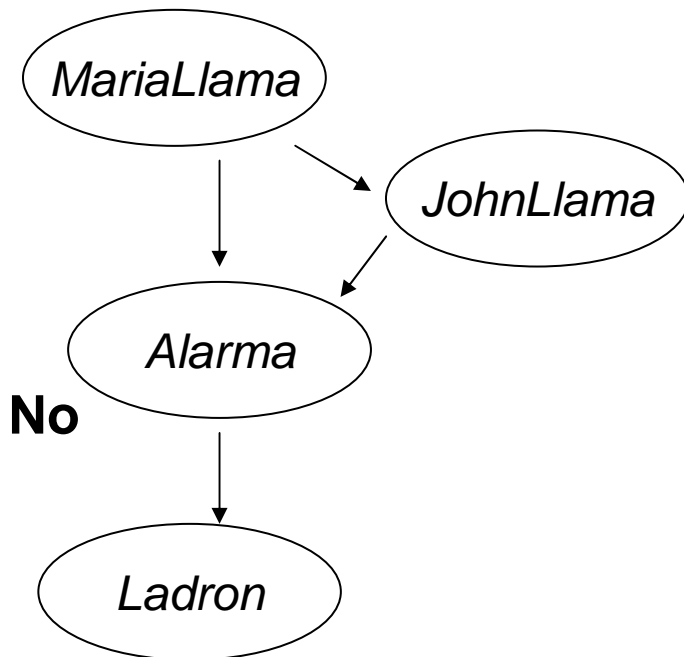
$P(B | A, J, M) = P(B | A)$?

$P(B | A, J, M) = P(B)$?

Redes Bayesianas : ¿Cómo construirlas?

Ejemplo

- Supongamos el siguiente orden M, J, A, B, E



$P(J | M) = P(J)$? **No**

$P(A | J, M) = P(A | J)$? $P(A | J, M) = P(A)$? **No**

$P(B | A, J, M) = P(B | A)$? **Si**

$P(B | A, J, M) = P(B)$? **No**

$P(E | B, A, J, M) = P(E | A)$?

$P(E | B, A, J, M) = P(E | A, B)$?

Redes Bayesianas : ¿Cómo construirlas?

Ejemplo

- Supongamos el siguiente orden M, J, A, B, E

$P(J | M) = P(J)$? **No**

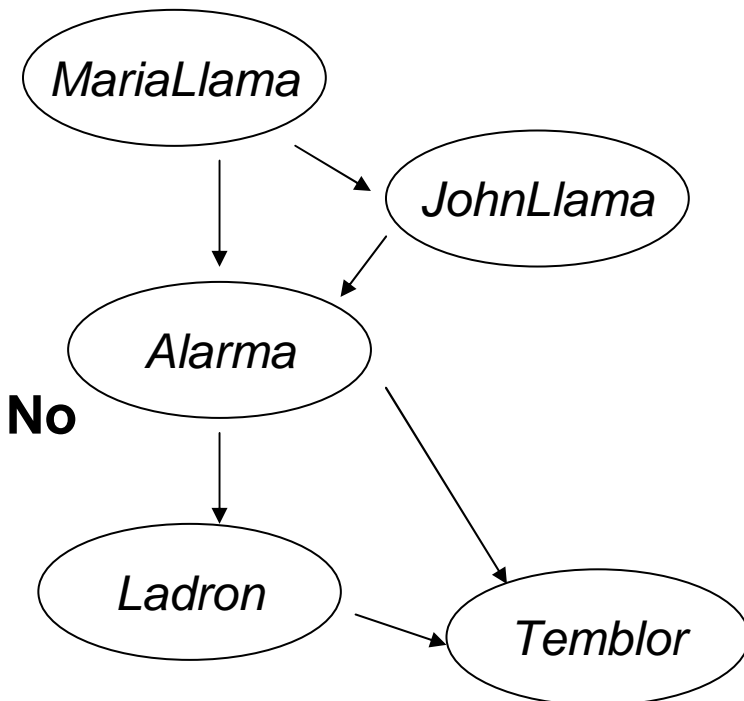
$P(A | J, M) = P(A | J)$? $P(A | J, M) = P(A)$? **No**

$P(B | A, J, M) = P(B | A)$? **Si**

$P(B | A, J, M) = P(B)$? **No**

$P(E | B, A, J, M) = P(E | A)$? **No**

$P(E | B, A, J, M) = P(E | A, B)$? **Si**



Redes Bayesianas : ¿Cómo construirlas?

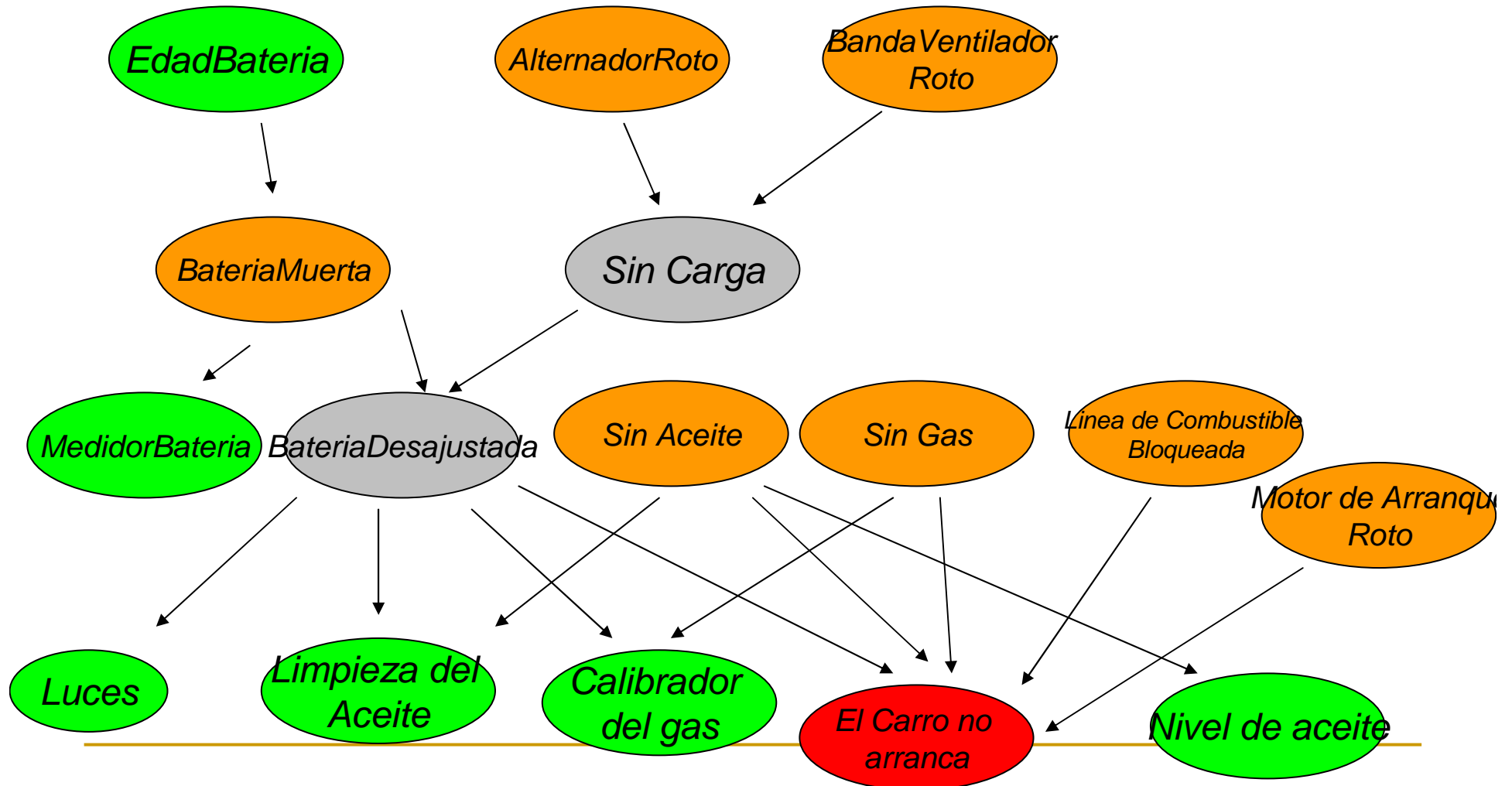
Ejemplo

- Ubicar independencia condicional es difícil en direcciones no causales
- Evaluar dichas dependencias es difícil en direcciones no causales
- La red es menos compacta $1 + 2 + 4 + 2 + 4 = 13$ números son necesarios

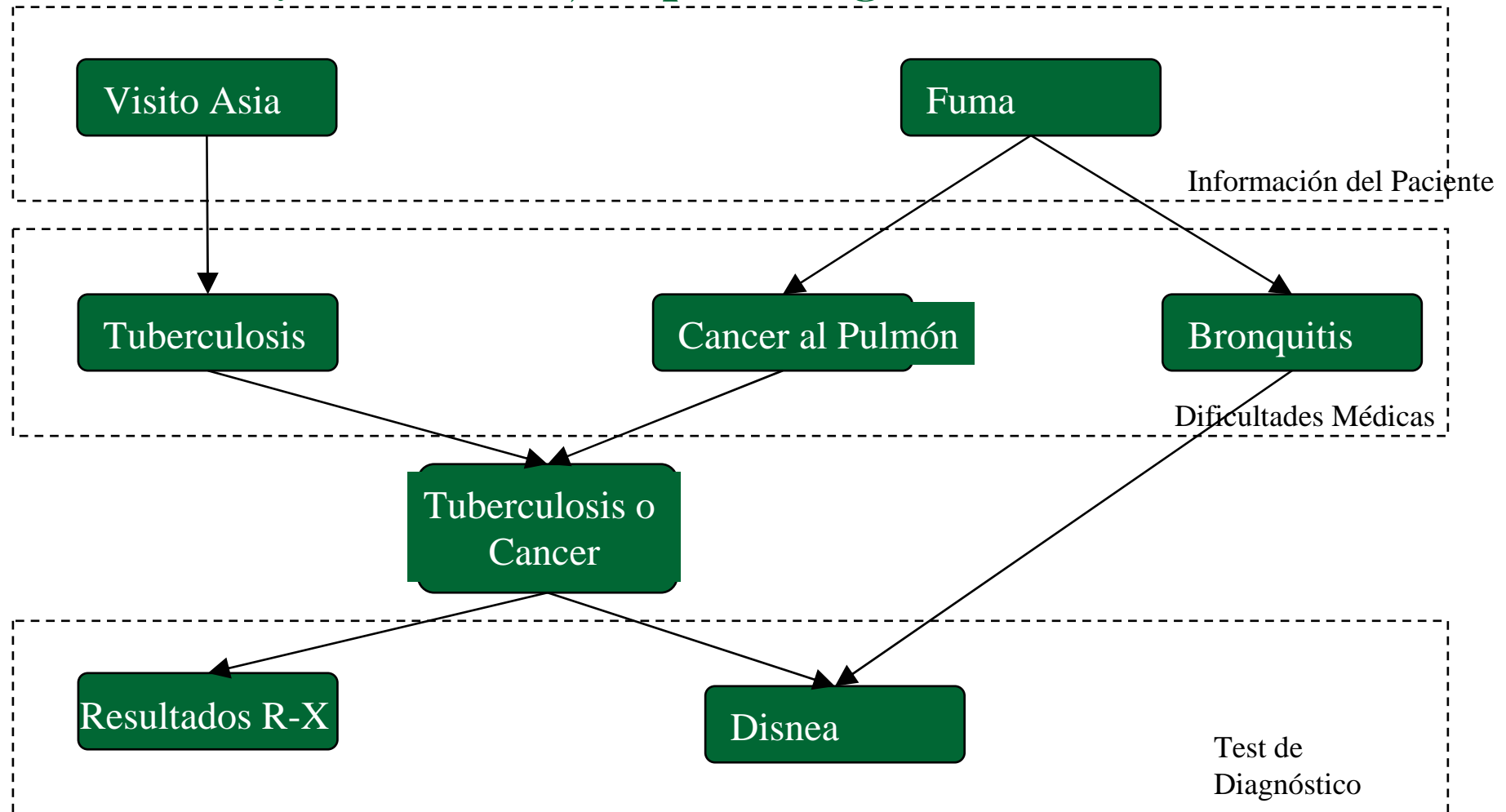
Redes Bayesianas : Ejemplo Diagnóstico de Carro

- Evidencia inicial : carro no arranca
- Variables verificables (verde)
- Variables “por lo que se puede arreglar el problema”(naranja)
- Variables ocultas (gris)

Redes Bayesianas : Ejemplo Diagnóstico de Carro



Redes Bayesianas : Ejemplo Diagnóstico Médico



- La red representa la estructura del conocimiento que modela la relación entre dificultades médicas, sus causas y efectos, información del paciente y test de diagnóstico

Referencias Bibliográficas

- Capitulo 14 Artificial Intelligence: A Modern Approach, Russell and Norvig
- <http://www.aaai.org/AITopics/pmwiki/pmwiki.php/AITopics/Uncertainty>