

# Tema 6. Redes Bayesianas

**Razonamiento y Representación del  
Conocimiento**

# Índice

- Teorema de Bayes (continuación)
- Redes Bayesianas
- Inferencia en redes Bayesianas
  - Inferencia exacta
  - Inferencia aproximada

# Teorema de Bayes (continuación)

- Repasemos el teorema de Bayes

$$P(Y/X) = \frac{P(X/Y)P(Y)}{P(X)} \qquad P(X) = \sum_i P(X/Y_i)P(Y_i)$$

- Concepto de independencia condicional

$$P(X, Y/Z) = P(X/Z)P(Y/Z) \qquad P(X/Y, Z) = P(X/Z)$$

- Los sucesos X e Y son condicionalmente independientes dado Z
- Regla de la cadena:

$$P(X, Y) = P(Y/X)P(X)$$

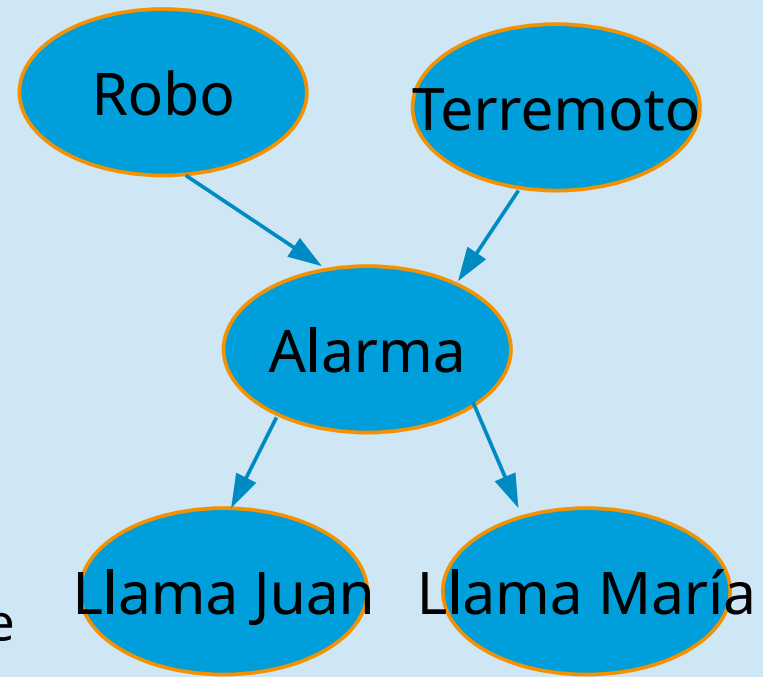
$$P(X, Y, Z) = P(Z/X, Y)P(Y, X) = P(Z/X, Y)P(Y/X)P(X)$$

# Redes Bayesianas

- Red Bayesiana: herramienta para representar conocimiento en un dominio incierto
  - Se representa por un grafo **dirigido acíclico**
  - Proporciona una representación de la distribución conjunta completa
- Está compuesta por:
  - Variables aleatorias → nodos del grafo
  - Dependencia entre variables → aristas del grafo

# Redes Bayesianas

- Semántica de las Redes Bayesianas
  - Ejemplo. Variables:
    - **Robo**: entran a robar en casa
    - **Terremoto**: se produce un terremoto
    - **Alarma**: suena la alarma
    - **Llama Juan/María**: mi vecino Juan/María me llama porque oye la alarma
  - ¿Qué significa esta red bayesiana?



# Redes Bayesianas

- Semántica de las Redes Bayesianas
  - ¿Qué probabilidad hay de que suene la alarma pero no se ha producido un robo ni un terremoto y tanto Juan como María llamen?

$$P(j \wedge m \wedge a \wedge \neg r \wedge \neg t)$$

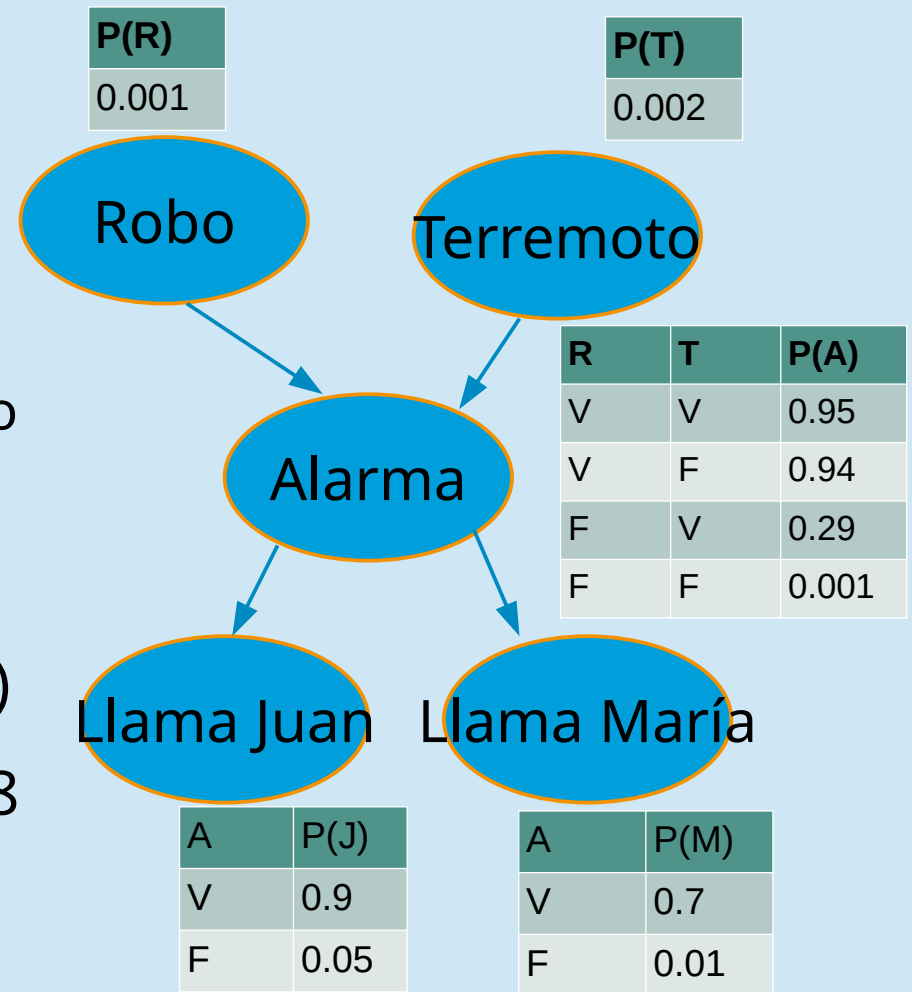
$$= P(j | a)P(m | a)P(a | \neg r, \neg t)P(\neg r)P(\neg t)$$

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

$$= 0.00062$$

- En general:

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

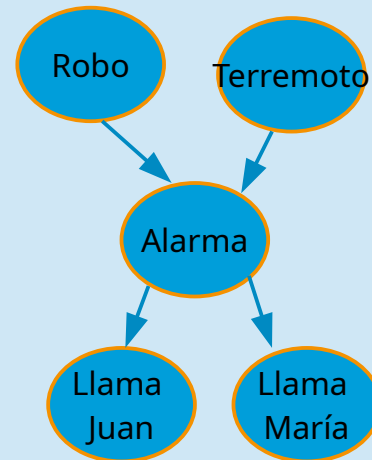


# Redes Bayesianas

- Construcción de Redes Bayesianas
  - Arbitraria: el diseñador decide las variables a considerar y sus relaciones de dependencia.
  - El diseño de la red afecta a su uso → coste temporal y espacial
  - La regla de diseño nos dice: el orden correcto en el que agregamos nodos es añadir las *causas raíces* primero, luego las variables que influyen, y así sucesivamente, hasta llegar a las hojas
  - Ejemplo: una RB con  $n$  ( $n=30$ ) variables influenciadas por al menos otras  $k$  ( $k=5$ ) puede especificarse por  $n \cdot 2^k$  valores (960). El mismo problema representado por una distribución de probabilidad conjunta requiere  $2^n$  valores (más de 1000 millones)

# Redes Bayesianas

- Independencia condicional
  - Al especificar una Red Bayesiana establecemos independencia condicional entre variables que no están directamente emparentadas en el grafo

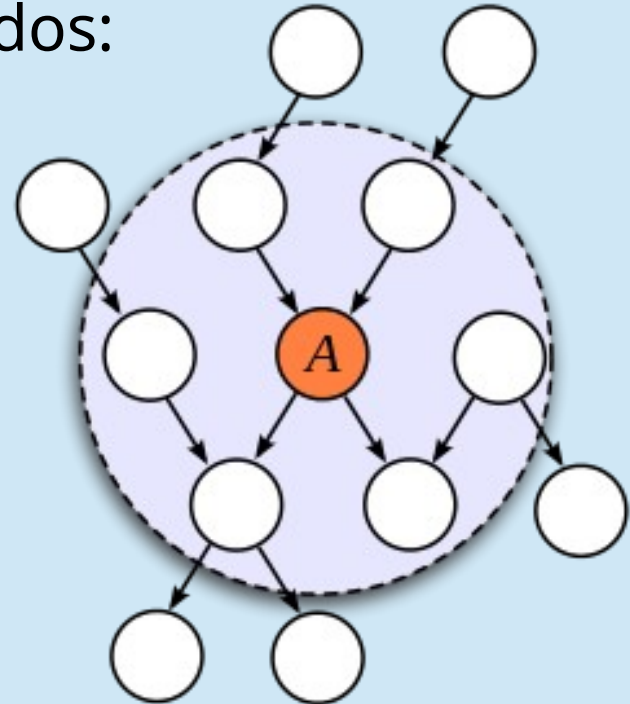


$$P(\text{LlamaMaría} \mid \text{LlamaJuan}, \text{Alarma}, \text{Terremoto}, \text{Robot}) = P(\text{LlamaMaría} \mid \text{Alarma})$$



# Redes Bayesianas

- Independencia Condicional
  - Cobertura de Markov:
    - Un nodo  $A$  es condicionalmente independiente de todos los nodos de la red dados:
      - Sus padres
      - Sus hijos
      - Los padres de sus hijos



# Inferencia en redes Bayesianas

- A partir de la distribución conjunta podemos contestar cualquier pregunta relativa a la red
- Varios tipos de inferencia en redes Bayesianas
  - Inferencia exacta (caso general)
  - Inferencia aproximada

# Inferencia en redes Bayesianas

- Inferencia exacta
  - Objetivo: computar la distribución de probabilidad a *posteriori* para un conjunto de **variables pregunta** dado algún evento observado (**variables evidencia**)
  - También consideramos las variables desconocidas (**variables ocultas**)
  - Inferencia por enumeración:

$$P(X/e) = \alpha P(X, e) = \alpha \sum_y P(X, e, y)$$

# Inferencia en redes Bayesianas

- Inferencia exacta (Ejemplo)
  - Pregunta:  $P(\text{Robo} \mid \text{LlamaJuan}=\text{cierto}, \text{LlamaMaría}=\text{cierto})$
  - **Variables pregunta:** Robo
  - **Variables evidencia:** LlamaJuan, LlamaMaría
  - **Variables ocultas:** Alarma, Terremoto
  - Inferencia por enumeración:

$$P(R/j, m) = \alpha P(R, j, m) = \alpha \sum_t \sum_a P(R, t, a, j, m)$$

Para Robo = cierto

$$P(R/j, m) = \alpha \sum_t \sum_a P(r) P(t) P(a/r, t) P(j/a) P(m/a)$$

# Inferencia en redes Bayesianas

- Inferencia exacta (Ejemplo)
  - Pregunta:  $\mathbf{P}(\text{Robo} \mid \text{LlamaJuan}=\text{cierto}, \text{LlamaMaría}=\text{cierto})$
  - Ejercicio 1: calcular para Robo = cierto

$$\mathbf{P}(R/j,m) = \alpha \sum_t \sum_a \mathbf{P}(r) P(t) P(a/r,t) P(j/a) P(m/a)$$

# Inferencia en redes Bayesianas

- Inferencia exacta (Ejemplo)
  - Pregunta:  $\mathbf{P}(\text{Robo} \mid \text{LlamaJuan}=\text{cierto}, \text{LlamaMaría}=\text{cierto})$
  - Ejercicio 1: calcular para Robo = cierto

$$\begin{aligned} \mathbf{P}(R/j,m) &= \alpha \sum_t \sum_a \mathbf{P}(r) P(t) P(a/r,t) P(j/a) P(m/a) \\ &= (0.001 * 0.002 * 0.95 * 0.9 * 0.7) + (0.001 * 0.002 * 0.05 * 0.05 * 0.01) + \\ &\quad (0.001 * 0.998 * 0.94 * 0.9 * 0.7) + (0.001 * 0.998 * 0.06 * 0.05 * 0.01) = \\ &\quad 0.00059224 \end{aligned}$$

# Inferencia en redes Bayesianas

- Inferencia exacta (Ejemplo)

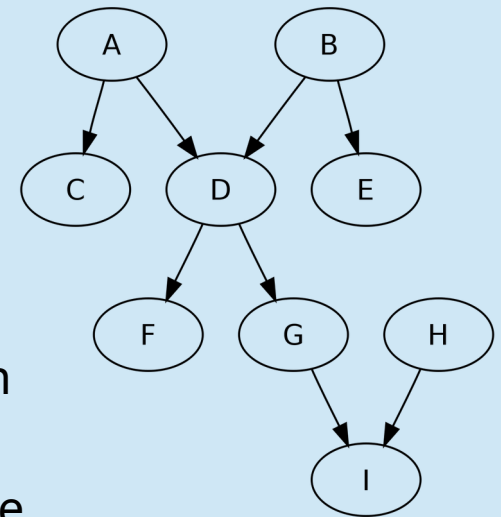
- Pregunta:  $P(\text{Robo} \mid \text{LlamaJuan}=\text{cierto}, \text{LlamaMaría}=\text{cierto})$
- Ejercicio 2: calcular para Robo = false

$$P(R/j,m) = \alpha \sum_t \sum_a P(\neg r) P(t) P(a/\neg r,t) P(j/a) P(m/a)$$

- Resultado: 0.0014919
- $P(R \mid j, m) = \alpha \langle 0.00059224, 0.0014919 \rangle = \langle 0.284, 0.716 \rangle$
- Ejercicio 3: calcular  $P(\text{LlamaJuan} \mid \text{Robo} = \text{cierto})$

# Inferencia en redes Bayesianas

- Inferencia exacta → Eficiencia
  - La inferencia exacta es muy costosa:  $O(2^n)$
  - Existen algoritmos más eficientes para tipos específicos de redes
    - Modelo de Kim & Pearl
      - Método de inferencia para redes bayesianas
      - Solo aplicable a un poliárbol: No existe más de un camino entre cada pareja de nodos
      - Se basa en el paso de dos tipos de mensajes entre nodos:
        - Para actualizar la credibilidad
        - Para introducir nueva evidencia
      - Se puede calcular en tiempo lineal  $O(n)$





# Inferencia en redes Bayesianas

- Inferencia aproximada
  - La inferencia exacta es enormemente costosa
  - La inferencia aproximada surge para dar solución
  - Se utilizan métodos de muestreo aleatorio (MonteCarlo)
    - Muestreo directo
    - Muestreo indirecto
    - Gibbs sampling

# Inferencia en redes Bayesianas

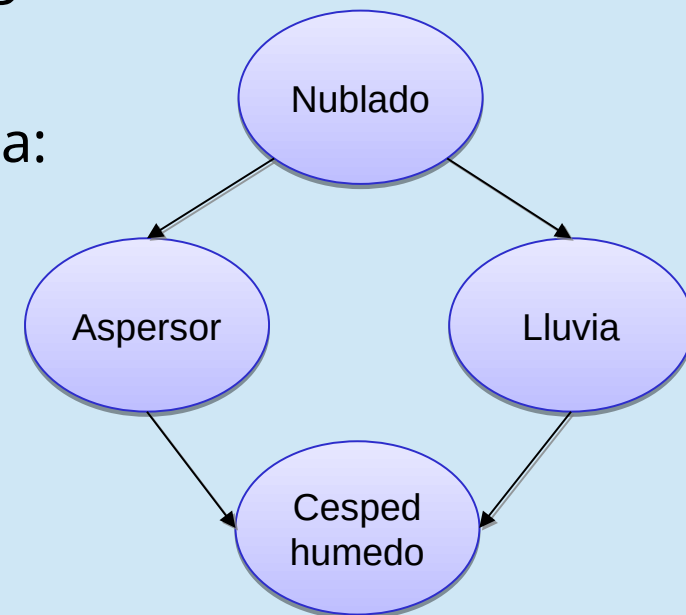
- Inferencia aproximada
  - Muestreo directo
    - Generar  $N$  muestras aleatoriamente a partir de los valores de la red
    - Contar cuántas muestras coinciden con la Pregunta que le hacemos a la red  $\rightarrow N_{MP}$
    - La probabilidad buscada aproximada =  $N_{MP}/N$
    - Coincide con la inferencia exacta cuando  $N \rightarrow \infty$

# Inferencia en redes Bayesianas

- Inferencia aproximada
  - Muestreo directo  $\rightarrow$  no tiene en cuenta la evidencia
  - Muestreo por rechazo
    - Generar  $N$  muestras aleatoriamente a partir de los valores de la red
    - Descartar aquellas muestras que no coinciden con la evidencia  $\rightarrow N_{\text{sel}}$
    - Contar cuántas muestras coinciden con la Pregunta que le hacemos a la red  $\rightarrow N_{\text{MP}}$
    - La probabilidad buscada aproximada =  $N_{\text{MP}}/N_{\text{sel}}$
    - Mejor aproximación de la probabilidad buscada

# Inferencia en redes Bayesianas

- Inferencia aproximada
  - Muestreo por rechazo - Ejemplo
    - Queremos estimar  $P(\text{Lluvia} \mid \text{Aspersor}=\text{cierto})$
    - Generamos 100 muestras de las que 73 tienen  $\text{aspersor}=\text{falso}$
    - De las 27 que coinciden con la evidencia:
      - En 8 Lluvia = cierto
      - En 19 Lluvia = falso
    - $P(\text{Lluvia} \mid \text{Aspersor}=\text{cierto}) = \text{Normalizar}(\langle 8, 19 \rangle) = \langle 0.296, 0.704 \rangle$



# Inferencia en redes Bayesianas

- Inferencia aproximada
  - Muestreo por rechazo → si hay mucha evidencia pueden rechazarse demasiadas muestras →  $N_{\text{sel}}$  bajo → mala aproximación
  - Muestreo por Gibbs o MCMC (MonteCarlo Markov Chain)
    - Las variables conocidas se fijan a su valor
    - Las variables ocultas se van muestreando siguiendo un orden arbitrario hasta obtener N muestras:
      - Cada variable oculta se muestrea dada su cobertura de markov

# Bibliografía recomendada

- Inteligencia Artificial. Un enfoque Moderno. Stuart Russell, Peter Noving. Ed. Prentice Hall. 2004
- <http://www.ecomportamiento.org/blog/2017/12/5/mejorando-nuestras-creencias-el-teorema-de-bayes>
- Pattern Recognition and Machine Learning. C. M. Bishop. Springer. 2006