Razonamiento con Incertidumbre

Lic. Carmen R. Garcia Perez

Probabilidades Condicionales

- Como el orden en que estén las variables en una **Función de Probabilidad Conjunta**, se puede escribir:

$$\mathrm{p}(\mathrm{Vi},\mathrm{Vj}) = p(Vi|Vj)\mathrm{p}(\mathrm{Vj}) = p(Vj|Vi)\mathrm{p}(\mathrm{Vi}) = \mathrm{p}(\mathrm{Vj},\mathrm{Vi})$$

- Entonces:

$$p(Vi|Vj) = \frac{p(Vj|Vi)\,\mathrm{p}(\mathrm{Vi})}{\mathrm{p}(\mathrm{Vj})}$$

- Esta última ecuación es muy importante. Se llama **Teorema de Bayes**.

,y
.

Inferencia Probabilística

- El escenario genérico para la Inferencia Probabilística es aquel en el que se tiene un conjunto V de variables proposicionales V_1, V_2, ….. V_k, en el que dan, en forma de evidencia, un subconjunto de variables de V, al que llamaremos E, con unos valores definidos E=e (Verdadero o Falso).

INFERENCIA PROBABILÍSTICA

- En las aplicaciones de agentes, las variables tendrán, por lo general, unos valores determinados. Lo que se desea es, calcular la probabilidad condicional :

$$p(Vi = vi | E = e)$$

- A este proceso se llamara **inferencia probabilística.**

,y
.

INFERENCIA PROBABILÍSTICA

- Como **Vi** puede tener los valores **Verdadero** o **Falso**, existen dos probabilidades condicionales las cuales son:

$$p(Vi = Verdadero | E = e)$$

Y

$$p(Vi = Falso | E = e)$$

INFERENCIA PROBABILÍSTICA

- Se va a utilizar un método de cálculo <<exhaustivo>> para obtener:

$$p(Vi = Verdadero | E = e)$$

- Si se utiliza la definición de **probabilidad condicional**, se tiene:

$$p(V_i = \text{Verdadero} \mid E{=}e) = \frac{p(Vi{=}Verdadero, E{=}e)}{p(E{=}e)}$$

INFERENCIA PROBABILÍSTICA

- El valor del numerador se obtiene utilizando la regla de cálculo de probabilidades de orden menor a partir de las de orden superior:

$$p(V_1 = Verdadero, \mathcal{E} = \mathbf{e}) = \sum_{V_i = Verdadero, \mathcal{E} = \mathbf{e}} p(V_1, \ \ldots V_k)$$

Donde : Vi, i=1,.., k define el conjunto de variables proposicionales. Es decir, se calcula la sumatoria de todos los valores de las probabilidades conjuntas de cada Vi=Verdadero, cuyos valores se obtienen de las variables de la evidencia. El cálculo de p(E=e) se puede realizar de forma similar.

INFERENCIA PROBABILÍSTICA

Como ejemplo, supongamos que tenemos las probabilidades conjuntas:

$$p(\mathrm{P}, \mathrm{Q}, \mathrm{R}) = 0{,}3$$
$$p(\mathrm{P}, \mathrm{Q}, \neg\mathrm{R}) = 0{,}2$$
$$p(\mathrm{P}, \neg\mathrm{Q}, \mathrm{R}) = 0{,}2$$
$$p(\mathrm{P}, \neg\mathrm{Q}, \neg\mathrm{R}) = 0{,}1$$
$$p(\neg\mathrm{P}, \mathrm{Q}, \mathrm{R}) = 0{,}05$$
$$p(\neg\mathrm{P}, \mathrm{Q}, \neg\mathrm{R}) = 0{,}1$$
$$p(\neg\mathrm{P}, \neg\mathrm{Q}, \mathrm{R}) = 0{,}05$$
$$p(\neg\mathrm{P}, \neg\mathrm{Q}, \neg\mathrm{R}) = 0{,}0$$

Se tiene $\neg\mathrm{R}$ como evidencia y se quiere calcular $p(\mathrm{Q} \mid \neg\mathrm{R})$

INFERENCIA PROBABILÍSTICA

Se tiene $\neg R$ como evidencia y se quiere calcular $p(Q \mid \neg R)$

$$p(Q \mid \neg R) = \frac{p(Q, \neg R)}{p(\neg R)} = \frac{[p(P, Q, \neg R) + p(\neg P, Q, \neg R)]}{p(\neg R)}$$

$$= \frac{(0,2 + 0,1)}{p(\neg R)} = \frac{0,3}{p(\neg R)}$$

$$p(\neg Q \mid \neg R) = \frac{p(\neg Q, \neg R)}{p(\neg R)} = \frac{[p(P, \neg Q, \neg R) + p(\neg P, \neg Q, \neg R)]}{p(\neg R)}$$

$$= \frac{(0,1 + 0,0)}{p(\neg R)} = \frac{0,1}{p(\neg R)}$$

$$p(P, Q, R) = 0,3$$
$$p(P, Q, \neg R) = 0,2$$
$$p(P, \neg Q, R) = 0,2$$
$$p(P, \neg Q, \neg R) = 0,1$$
$$p(\neg P, Q, R) = 0,05$$
$$p(\neg P, Q, \neg R) = 0,1$$
$$p(\neg P, \neg Q, R) = 0,05$$
$$p(\neg P, \neg Q, \neg R) = 0,0$$

Se sabe que : $p(Q \mid \neg R) + p(\neg Q \mid \neg R) = 1$, entonces se tiene:

$$p(Q \mid \neg R) = 0,75$$

INFERENCIA PROBABILÍSTICA

La Inferencia probabilística mediante este método es intratable, ya que cuando se tiene k variables se necesita una lista de 2^k probabilidades conjuntas $p(V_1, V_2, \ldots\ldots, V_k)$.

Los cálculos de las probabilidades condicionales de ciertas variables, dada la evidencia, sobre ellas. La forma de representar el conocimiento eficientemente involucra lo que se denominan **Independencias Condicionales.**

INDEPENDENCIA CONDICIONAL

- Una variable V es **condicionalmente independiente** de un conjunto de variables V_i, dado otro conjunto V_j, cuando:

$$p(V|V_i,V_j) = p(V|V_j).$$

- Como generalización de la independencia entre parejas, se dice que las variables $V_1,\ldots V_k$ son mutuamente independientes, dado un conjunto V, si cada una de las variables es condicionalmente independiente de todas las demás, dado V como:

INDEPENDENCIA CONDICIONAL

$$p(V_1, V_2, \ldots, V\kappa | V) = \prod_{i=1}^{k} p(V_i | V_{i-1}, V_2, \ldots, V_1, V)$$

- y como cada V_i es **condicionalmente independiente** de las otras dado V, se tiene que:

$$p(V_1, V_2, \ldots, V\kappa | V) = \prod_{i=1}^{k} p(V_i | V)$$

- En el caso de que V esté vacío, se tiene:

$$p(V_1, V_2, \ldots, V\kappa) = p(V_1)p(V_2)..p(Vk)$$

- Se dice que las variables son incondicionalmente independientes.

INDEPENDENCIA CONDICIONAL

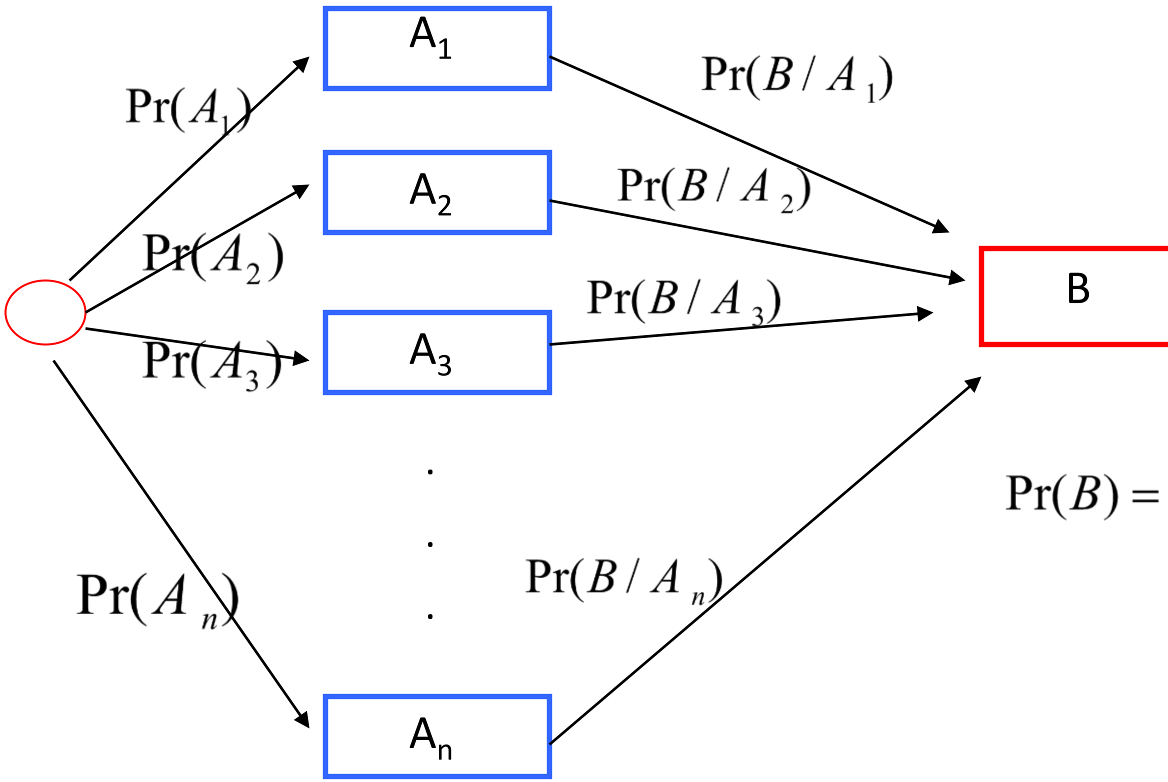
- Las independencias condicionales se pueden representar adecuadamente con las estructuras denominadas **Redes Bayesianas (**También se las denomina **Redes de Creencias)**.
- Estas estructuras son muy útiles en la inferencia probabilística.
- Las independencias condicionales representadas por Redes Bayesianas llevan a grandes ahorros en los cálculos de Inferencia Probabilística.

Teorema de Bayes

- Supongamos que sobre el espacio muestral S se tiene una partición A_i, con i = 1, ..., n.

- Esto significa que cualquier resultado de S necesariamente debe estar en uno y solo uno de los eventos A_i

Teorema de Bayes

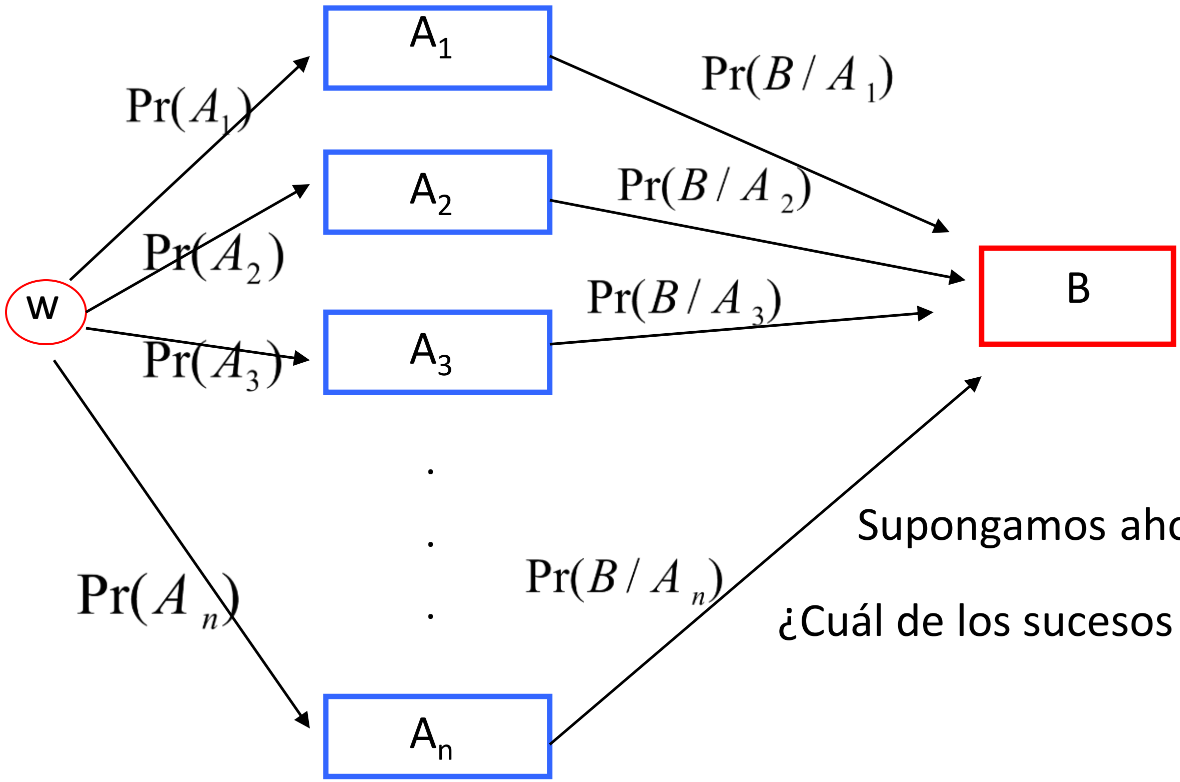


$$\Pr(B) = \sum_{i=1}^{n} \Pr(B / A_i)\Pr(A_i)$$

Este cálculo es para medir la incertidumbre de la ocurrencia del evento B.

Medición del futuro, representado por el evento B

Teorema de Bayes

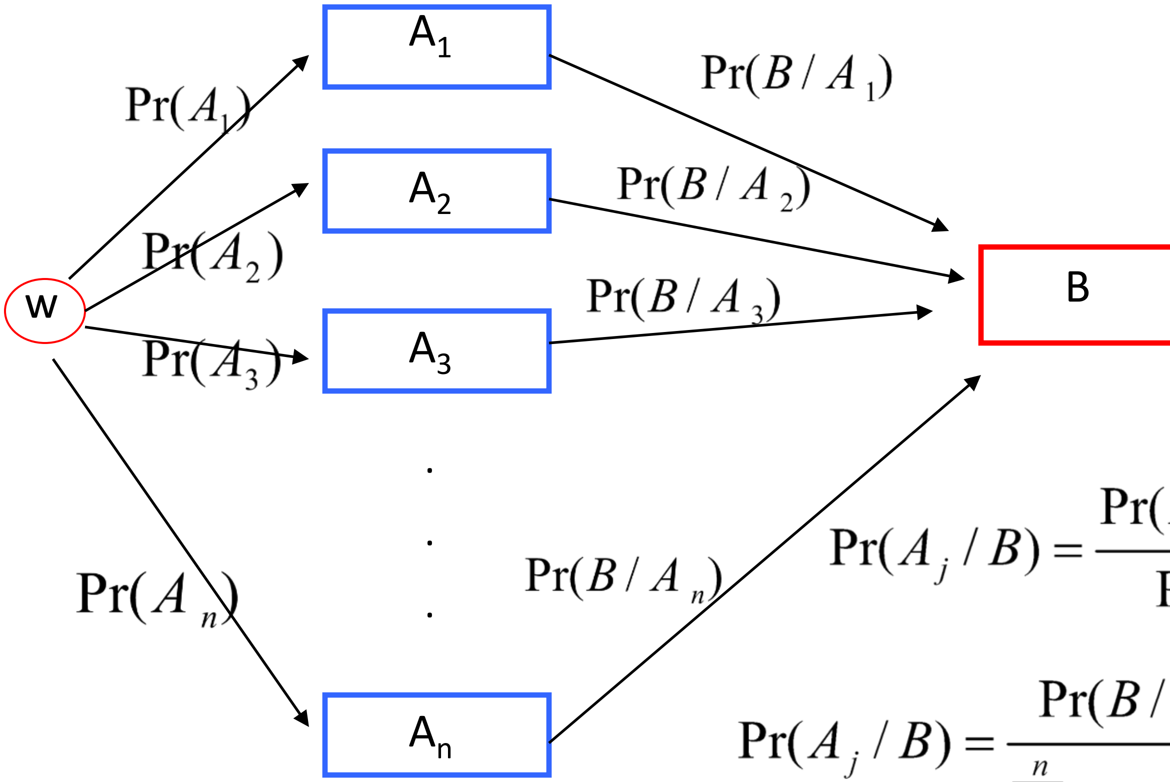


Supongamos ahora que B ocurre.

¿Cuál de los sucesos A_j ha ocurrido?

De otra forma, ¿cuál es el valor de $\Pr(A_j / B)$ con j = 1, ...n?

Teorema de Bayes



$$\Pr(A_j / B) = \frac{\Pr(A_j \cap B)}{\Pr(B)} = \frac{\Pr(B / A_j) \cdot \Pr(A_j)}{\Pr(B)}$$

$$\Pr(A_j / B) = \frac{\Pr(B / A_j) \cdot \Pr(A_j)}{\sum_{i=1}^{n} P(B / A_i) \cdot \Pr(A_i)}$$

Medición del pasado, representado por el evento A_j

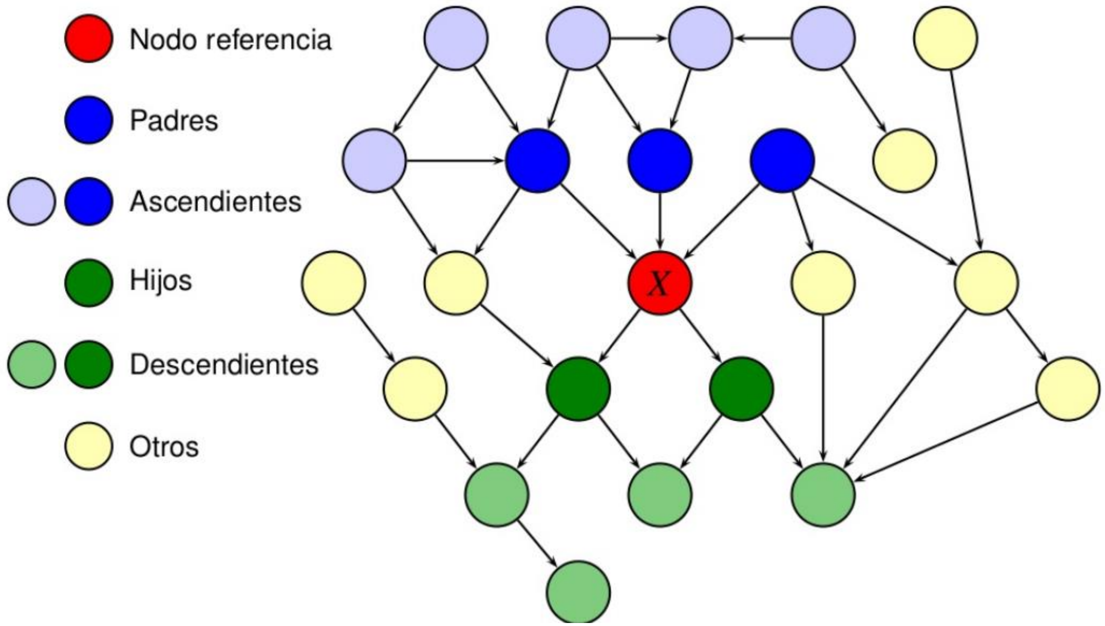
Redes Bayesianas

- Una **red bayesiana** es un grafo dirigido acíclico cuyos nodos están etiquetados con variables aleatorias.

- Una **red bayesiana** estipula que cada nodo V_i del grafo es condicionalmente independiente de cualquier subconjunto de nodos que no sean descendientes de V_i, dados sus padres.

Redes Bayesianas

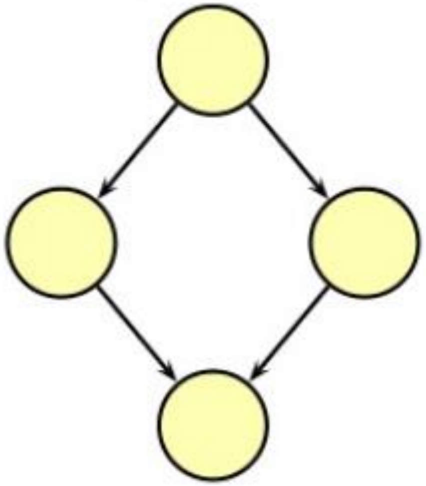
- La red bayesiana es una gráfica en la que se cumple lo siguiente:
  - Los nodos de la red están formados por un conjunto de variables aleatorias.
  - Cada par de nodos se conecta entre sí mediante un conjunto de enlaces o flechas. El significado implícito de una flecha que vaya del nodo *X* al nodo *Y* es el de que *X* ejerce una *influencia directa* sobre *Y*.

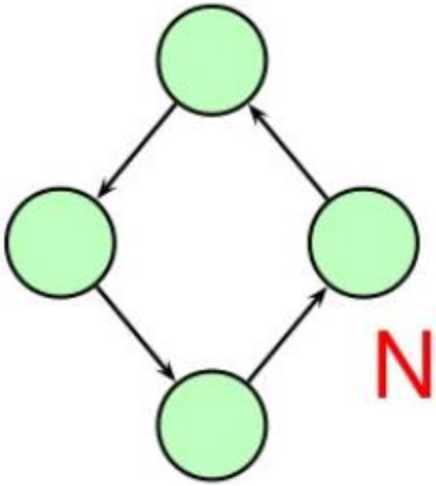


Redes Bayesianas

- La red bayesiana es una gráfica en la que se cumple lo siguiente:
  - Por cada nodo hay una tabla de probabilidad condicional que sirve para cuantificar los efectos de los padres sobre el nodo. Los padres de un nodo son aquellos cuyas flechas apuntan hacia al nodo.
  - La gráfica no tiene ciclos dirigidos (es acíclica).

SI





NO

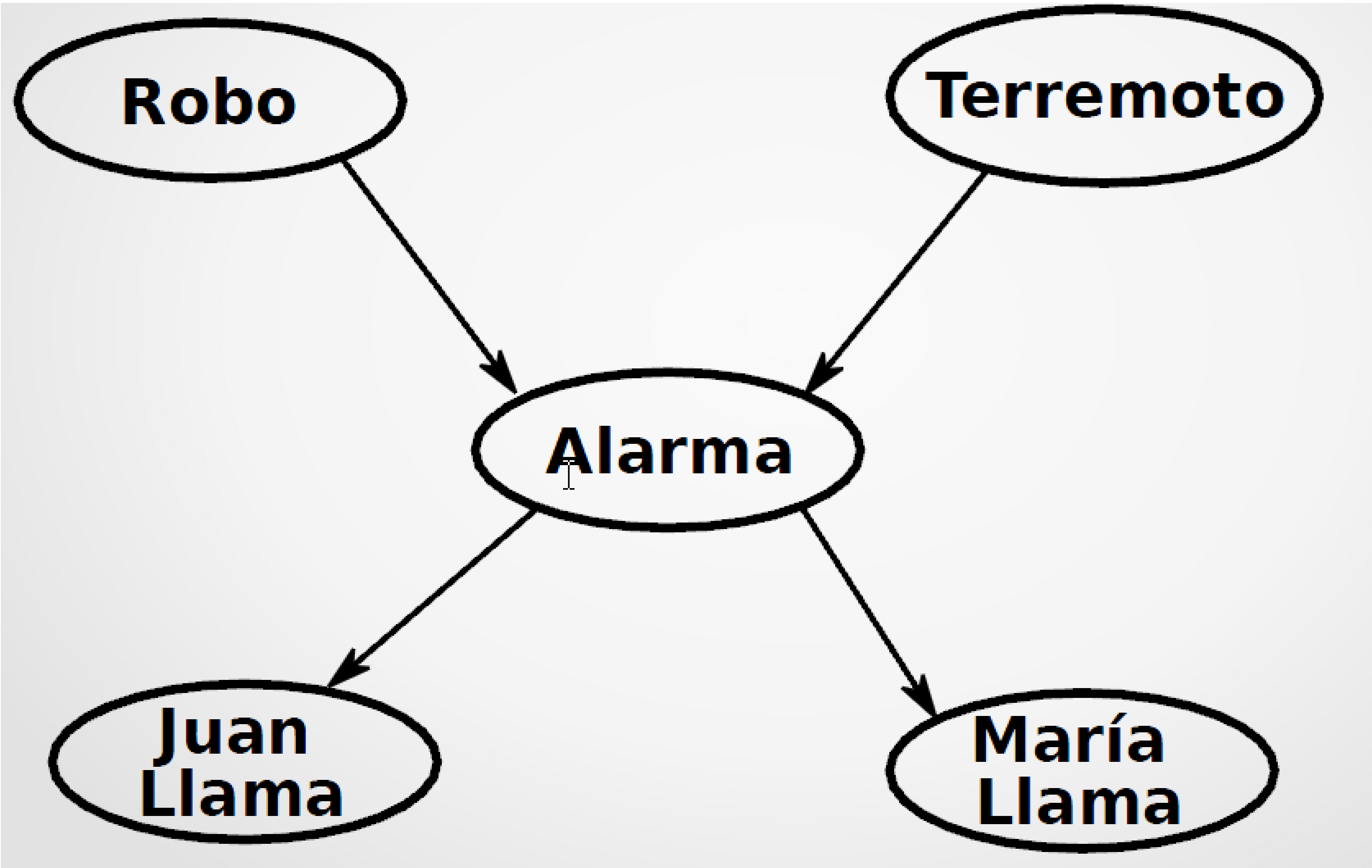
Redes Bayesianas

- Sean V_1, V_2,…,V_k los nodos de una **red bayesiana.** Con la independencia condicional que se ha descrito para la red, se puede definir la probabilidad conjunta de todos los nodos como:

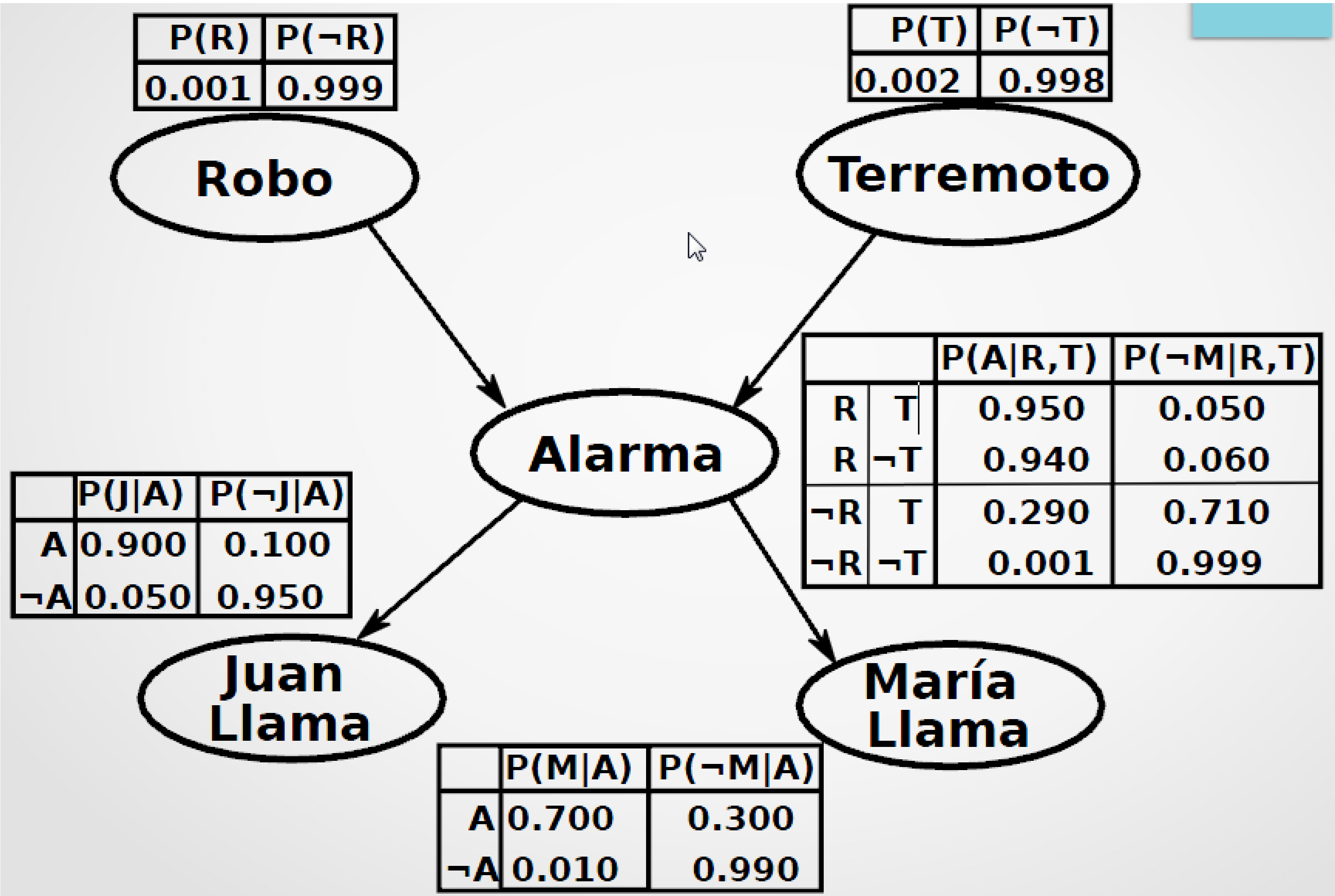
$$p(V_1, V_2 \ldots, V_k) = \prod_{i=1}^{k} p(V_i \,|P(V_i\,))$$

Donde: $P(V_i\,)$ son los padres de V_i en el grafo.

Ejemplo 1: Redes Bayesianas



Ejemplo 1: Redes Bayesianas



P(R)	P(¬R)
0.001	0.999

P(T)	P(¬T)
0.002	0.998

		P(A|R,T)	P(¬M|R,T)
R	T	0.950	0.050
R	¬T	0.940	0.060
¬R	T	0.290	0.710
¬R	¬T	0.001	0.999

	P(J|A)	P(¬J|A)
A	0.900	0.100
¬A	0.050	0.950

	P(M|A)	P(¬M|A)
A	0.700	0.300
¬A	0.010	0.990

Ejemplo 1: Redes Bayesianas

Probabilidad Conjunta

$P(j,m,a,\sim r,\sim t)= P(j|m,a,\sim r,\sim t)P(m,a,\sim r,\sim t)= P(j|m,a,\sim r,\sim t)P(m|a,\sim r,\sim t)P(a,\sim r,\sim t)= P(j|m,a,\sim r,\sim t)P(m|a,\sim r,\sim t)P(a|\sim r,\sim t)P(\sim r,\sim t)= P(j|m,a,\sim r,\sim t)P(m|a,\sim r,\sim t)P(a|\sim r,\sim t)P(\sim r|\sim t)P(\sim t)$ (Este cálculo con probabilidades).

Usando con Redes Bayesianas:

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

$P(j,m,a,\sim r,\sim t)=0.9*0.7*0.001*0.999*0.998= 0.000628$

Ejemplo 2: Redes Bayesianas

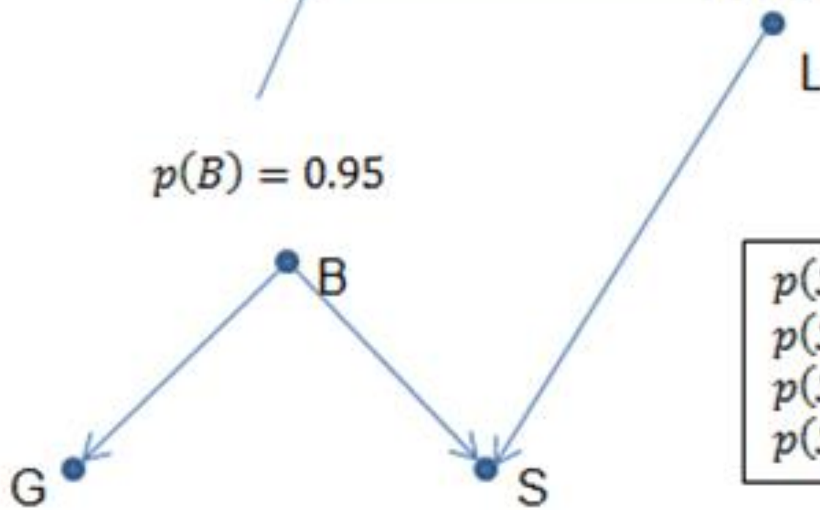
- La construcción de una **red bayesiana** mediante el ejemplo del apilamiento de bloques. Se empieza con las primeras causas:
  - <la Batería está cargada> (B)
  - <el bloque se puede Levantar> (L)
  - B y L son causas de S (<el brazo se mueve>)
  - B es una causa de G (<el indicador muestra que la batería está cargada>).
- Se tiene la siguiente **Red Bayesiana**:

Redes Bayesianas

Probabilidades a priori asociadas a cada nodo sin padres

$p(L) = 0.7$

$p(B) = 0.95$



$p(S|\ B, L) = 0{,}9$
$p(S|\ B, \neg L) = 0{,}05$
$p(S|\neg\ B, L) = 0{,}0$
$p(S|\ \neg B, \neg L) = 0{,}0$

$p(G|B) = 0.95$
$p(G|\neg\ B) = 0.1$

Tablas de probabilidades condicionales asociadas a cada nodo con padres

$$p(G, B, S, L) = \mathrm{p(G|B)p(S|B,L)p(B)p(L)}$$

PATRONES DE INFERENCIA EN REDES BAYESIANAS

- Hay tres patrones importantes de inferencia en las redes bayesianas.

1. **Inferencia Causal (o Descendente)**

(de las causas a los efectos).

2. **Inferencia de Diagnóstico (Ascendente)**

(de los efectos a las causas).

3. **Inferencia Intercausal (Justificación)**

(entre las causas de un efecto común).