Tema 1 "Sistemas basados en el conocimiento"

1. Introducción

Usa una base de conocimiento para razonar y resolver problemas complejos.

Funcionalidades ⇒
- * Aceptar las consultas
- * Aceptar los datos y solicitar otros
- * Procesar la información
- * Emitir y justificar la respuesta hallada
- * Autoaprendizaje

2. Arquitectura y motores de inferencia

¡SBC basado en reglas!

- **Base de conocimientos**

Contiene el conocimiento experto → relaciones entre los objetos

Naturaleza estática → no cambia a menos que se introduzcan nuevas reglas

⇓

$\boxed{\text{Si ... entonces ...}}$

- **Memoria de trabajo**

Contiene las conclusiones o hipótesis que se generan en el proceso de razonamiento a partir de los datos del caso.

Naturaleza dinámica ⟶ cambia de una sesión a otra

- **Motores de inferencia**

Datos del caso
+
Datos memoria de trabajo ⟹ obtiene nuevas conclusiones o hipótesis
+
Conocimiento de la BC

(se añaden a la memoria de trabajo)

Puede solicitar nuevos datos al usuario

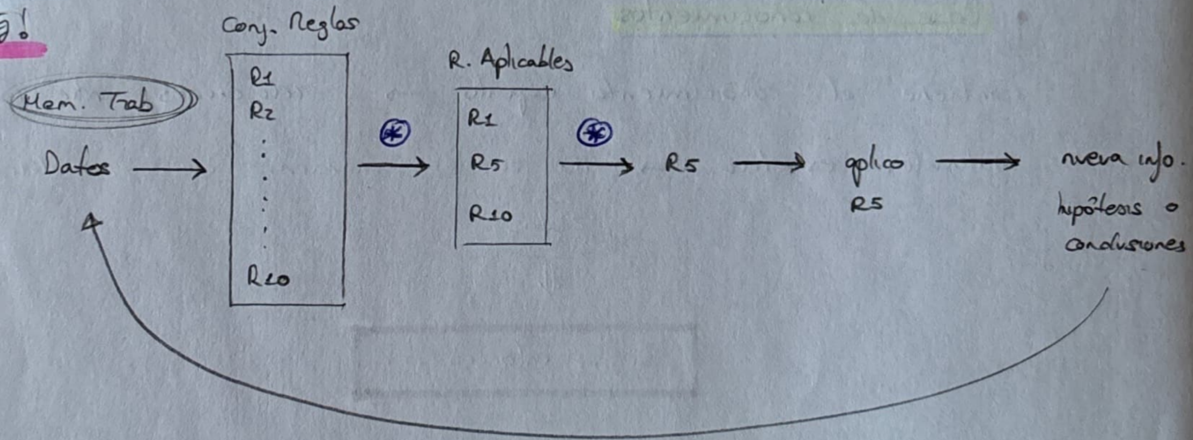
MODUS PONENS

$$\frac{p \rightarrow q \qquad p}{q}$$

MODUS TOLLENS

$$\frac{p \rightarrow q \qquad \neg q}{\neg p}$$

¡Ej!



(*) Estrategias

(*) Resolver conflictos

→ Estrategias de los motores de inferencia:

① Encadenamiento de reglas hacia delante (partimos de hechos conocidos)

② Encadenamiento de reglas hacia atrás (partimos de un objetivo a demostrar)

¡Merda!

➡ Criterios de resolución de conflictos: estáticos, dinámicos o dinámicos manipulables. Ej. → orden textual

➡ Criterios de parada Ej. → no es posible aplicar más reglas

Tema 2 "Redes bayesianas"

1. Conceptos básicos de probabilidad

- **Probabilidad condicionada**

$$\boxed{P(x/y) = \frac{P(x,y)}{P(y)}}$$

Si x e y son independientes $\Rightarrow$ $P(x,y) = P(x) \cdot P(y)$

Si x e y son dependientes $\Rightarrow$ $\begin{cases} P(x,y) = P(x) \cdot P(y/x) \\ P(x,y) = P(y) \cdot P(x/y) \end{cases}$

- **Teorema de Bayes**

$$\boxed{P(y/x) = \frac{P(y) \cdot P(x/y)}{P(x)}}$$

- **Ley de probabilidad total**

$$P(x) = P(y_1) \cdot P(x/y_1) + P(y_2) \cdot P(x/y_2) + \dots$$

- **Probabilidad marginal a partir de la distribución conjunta**

$$P(X_j = x_j) = \sum P(x_1, x_2, \dots, x_n)$$

- **Independencia condicional** $P(X, Y / Z) = P(X/Z) \cdot P(Y/Z)$

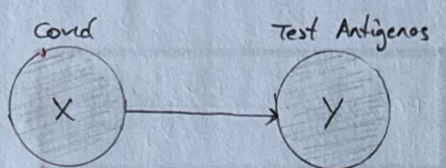
2. Presentación intuitiva

Nodo → variable o entidad del mundo real (X)

Valor → x ($+x$ = sí , $\neg x$ = no)

Arco → une dos nodos, representa una relación de influencia causal

¡Ej!



$p(+x) = 0{,}099$ (tengo covid)

$p(+y/+x) = 0{,}9$

$p(+y/\neg x) = 0{,}05$

Sacamos:

- Probabilidad no tengo covid $\Rightarrow P(\neg x) = 1 - 0{,}099 = \boxed{0{,}901}$
- Probabilidad tengo covid y no doy positivo $\Rightarrow P(\neg y/+x) = \boxed{0{,}1}$
- Probabilidad no tengo covid y no doy positivo $\Rightarrow P(\neg y/\neg x) = \boxed{0{,}95}$

CONCEPTO DE INFERENCIA

- Probabilidad a priori de que una persona de positivo en un test:

$$P(+y) = P(+x) \cdot P(+y/+x) + P(\neg x) \cdot P(+y/\neg x) = \boxed{0{,}134}$$

$$P(\neg y) = 1 - P(+y) = \boxed{0{,}865}$$

Inferencia a partir de datos $\Rightarrow$ probabilidad a posteriori dada una evidencia observada

* Si he dado positivo en el test, ¿cuál es la probabilidad de que tenga cond?

$$\underline{P^*(x) = P(x/e)} \longrightarrow \text{el hecho que sé}$$

$$P^*(+x) = P(+x/+y) = \frac{P(+x) \cdot P(+y/+x)}{P(+y)} = \boxed{0,664}$$

¡Ojo! $\Longrightarrow$ expresión general del teorema de Bayes

$$P^*(x) = P(x/y) = \frac{P(x) \cdot P(y/x)}{P(y)} \Longrightarrow P^*(x) = \underbrace{\alpha}_{\frac{1}{P(y)}} \cdot P(x) \cdot \underbrace{\lambda(x)}_{P(y/x)}$$

Sacamos α normalizando $(\square + \square = 1)$

obtenemos $P^*(x)$

3. Definición formal de R. Bayesiana (Ejercicio Tipo)

Formada por:

Variables proposicionales $\rightarrow$ variable aleatoria que toma un conjunto exhaustivo y excluyente de valores

Variables proposicionales + Conjunto de relaciones binarias + Distribución de probabilidad conjunta

Cumple que:

$(V, E) \longrightarrow$ grafo acíclico, conexo y dirigido

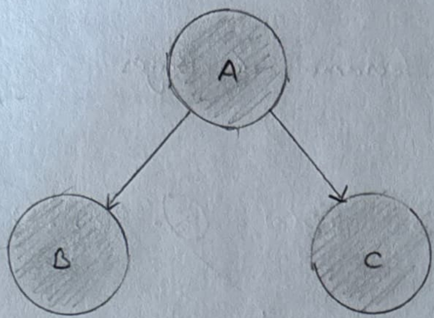
$(C, P) \longrightarrow$ hipótesis de independencia condicional

- **Hipótesis independencia condicional**

$\forall x \in V$ y $\forall y \in V - \{x \cup \text{descendientes}(x) \cup \text{padres}(x)\}$

$\boxed{x \text{ es independiente de } y \text{ dado padres}(x)}$

¡Ej Tipo! Comprobar si una red es bayesiana



- $x = A \Rightarrow A$; $y \in \emptyset$
- $x = B \Rightarrow B$; $y \in \{C\}$ dado A
- $x = C \Rightarrow C$; $y \in \{B\}$ dado A

Comprobar si B es independiente de C dado A ¿$P(b_1/a_1, c_1) \overset{?}{=} P(b_1/c_1)$

$$P(b_1/a_1, c_1) = \frac{P(a_1, b_1, c_1)}{P(a_1, c_1)} = \frac{1}{4}$$

$\boxed{\text{NO}}$

$$P(b_1/c_1) = \frac{P(b_1, c_1)}{P(c_1)} = \frac{3}{7}$$

4. Teorema Fundamental. Factorización de la probabilidad

La distribución de probabilidad de una red bayesiana puede expresarse:

$$\boxed{P(x_1, \dots, x_n) = \prod P(x_i / \text{padre}(x))}$$

sino se requieren muchas prob. conjuntas ¡!

5. Modelado con redes bayesianas

Modelado
- → cualitativo: nodos o variables, relaciones
- → cuantitativo:
  - prob. a priori nodos sin padre
  - prob. condicionada de cada nodo dado su padre

IMPORTANTE

- Quiero representar las enfermedades ⇒ mejor nodos separados con valores si/no porque puedo tener varias enfermedades a la vez

- Si los valores de una variable son únicos (solo puede toma uno) ⇒ creo una variable con todos los casos ¡Ej! especie

- Relaciones de influencia causal ⇒ "si gripe entonces tos"

  gripe → tos   pájaro → vuela

- Las probabilidades deben ser acordes con el sentido común

- Diferencia entre el OR y AND ⇒
  true && true = true
  true && false = false
  true || false = true
  (...)

- Pasos para modelar una red bayesiana

1º Identificar información relevante del problema

2º Representar variables y conjunto de valores ¡Excluyente y exhaustivo!

3º Identificar relaciones de influencia causal

4º Obtener las probabilidades necesarias ¡Coherentes!

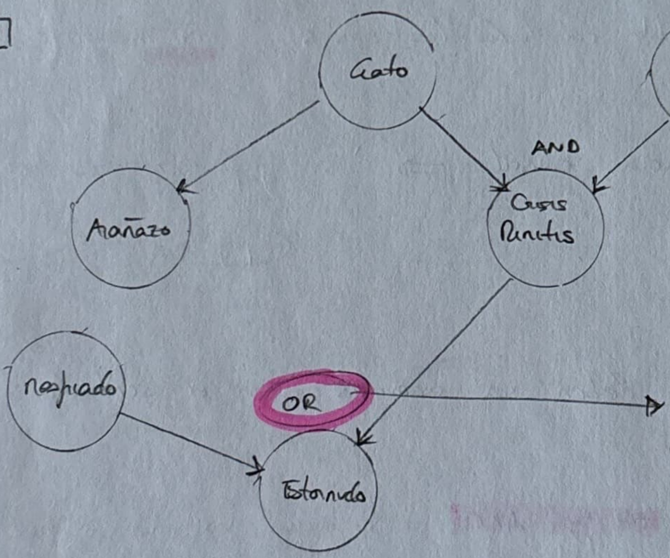
¡5'1!

$\boxed{1,2}$ Estornuda = {sí, no} Gato = {sí, no}

Resfriado = {sí, no} Rinitis = {sí, no}

Arañado = {sí, no} Alergia = {sí, no}

$\boxed{3}$



el estornudo puede ser causado por un resfriado o por la rinitis

$\boxed{4}$ $P(G) = 0{,}2$ $P(A) = 0{,}45$ $P(Res) = 0{,}5$

$P(\text{arañazo} / \neg \text{gato}) = 0{,}15$ $P(\text{arañazo} / \text{gato}) = 0{,}95$

$P(\text{rinitis} / \text{gato}, \text{alergia}) = 0{,}99$

$P(\text{rinitis} / \neg \text{gato}, \text{alergia}) = 0{,}1$

$P(\text{rinitis} / \text{gato}, \neg \text{alergia}) = 0{,}15$

$P(\text{rinitis} / \neg \text{gato}, \neg \text{alergia}) = 0{,}005$

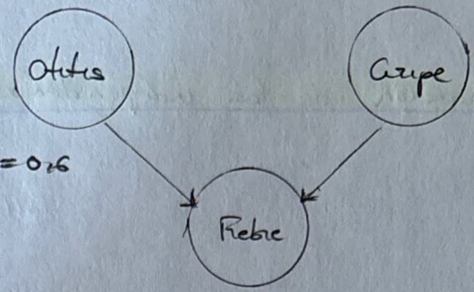
$P(\text{estornudo} / \text{resfriado}, \text{rinitis}) = 0{,}99$

$P(\text{estornudo} / \neg \text{resfriado}, \text{rinitis}) = 0{,}9$

$P(\text{estornudo} / \text{resfriado}, \neg \text{rinitis}) = 0{,}85$

$P(\text{estornudo} / \neg \text{resfriado}, \neg \text{rinitis}) = 0{,}01$

Ej.



$c_o = P(+f/+o) = 0,6$ $\qquad$ $c_g = P(+f/+g) = 0,8$

- Puerta OR $\quad \boxed{\text{objetivo}} \longrightarrow P(F/O,G)$

$P(+f/\neg g, \neg o) = 0$

$P(+f/+g, \neg o) = P(+f/+g) + P(\neg f/+g) \cdot \cancel{P(+f/\neg g, \neg o)}^{\,0} = c_g = 0,8$

$P(+f/\neg g, +o) = P(+f/+o) + P(\neg f/+o) \cdot \cancel{P(+f/\neg g, \neg o)}^{\,0} = c_o = 0,6$

$P(+f/+g, +o) = P(+f/+g) + P(\neg f/+g) \cdot P(+f/\neg g, +o) =$

$= 0,8 + (0,2 \cdot 0,6) = \boxed{0,92}$

- Puerta Noisy - or

$P(+f/\neg g, \neg o) = c_z = 0,01 \quad$ (ruido)

H1, H4 $\Longrightarrow P(+f/+g, +o, \neg z) = P(+f/+g, \neg o, \neg z) + P(\neg f/+g, \neg o, \neg z) \cdot$

$P(+f/\neg g, +o, \neg z) = 0,8 + (0,2 \cdot 0,6) = \boxed{0,92}$

H2 $\Longrightarrow P(+f/\neg g, \neg o, \neg z) = 0$

$P(+f/+g, +o) = P(+f/+g, +o, \neg z) + P(\neg f/+g, +o, \neg z) \cdot P(+f/\neg g, \neg o) =$

$= 0,92 + (0,08 \cdot 0,01) = \boxed{0,9208}$

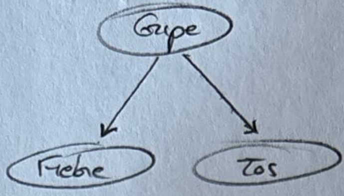
$P(+f/+g, \neg o) = P(+f/+g, \neg o, \neg z) + P(\neg f/+g, \neg o, \neg z) \cdot P(+f/\neg g, \neg o) =$

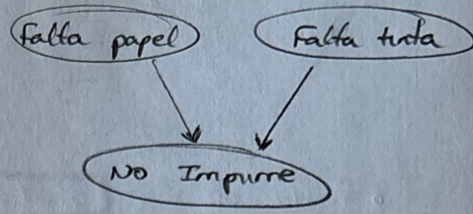
$= 0,8 + (0,2 \cdot 0,01) = \boxed{0,802}$

$P(+f/\neg g, +o) = P(+f/\neg g, +o, \neg z) + P(\neg f/\neg g, +o, \neg z) \cdot P(+f/\neg g, \neg o) =$

$= 0,6 + (0,4 \cdot 0,01) = \boxed{0,604}$

- **Variables: verificación de las independencias**





Fiebre y Tos son <u>dependientes</u> a priori pero <u>independientes</u> dado gripe

- **Parámetros** (probabilidades)

→ Especificación directa (and, or) : normalmente si no tengo info. plena, asumo equiprobabilidad

Modelo Noisy-OR

H1 * Cada causa, por sí misma, puede producir el efecto

H2 * <u>Ruido</u> → hay causas desconocidas, no incluidas en la red, que pueden producir el efecto

H3 * No hay interacción entre las causas

H4 * El efecto se produce solo por alguna causa modelada o por ruido.

$\Downarrow$

Dada las probs. cond. por un solo padre y el factor ruido, se pueden calcular las probs. cond. por todos los padres

Modelo Noisy-And

$\neg(A \text{ and } B) \equiv \neg A \text{ or } \neg B$

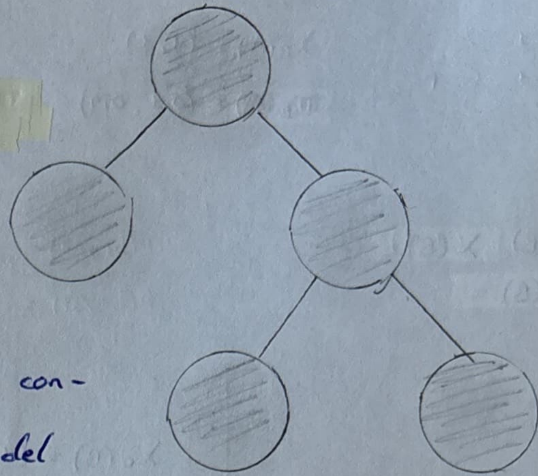
$\neg(A \text{ or } B) \equiv \neg A \text{ and } \neg B$

$\Rightarrow$ $\boxed{\text{Truco}}$ $P(\neg x / u_1, \dots, u_n) = \prod_{i \in T_u} q_i$

(conjunto de causas de x presentes)

6. Algoritmo de propagación de probabilidades en árboles

Árbol (cada nodo a lo sumo un padre)



Algoritmo → sirve para calcular las nuevas probs de los nodos conforme se va teniendo evidencia del caso concreto

- **Idea general**

$$\boxed{P(x) = \alpha \cdot \lambda(x) \cdot \pi(x)}$$

→ λ - valor (cambia cuando un hijo de x se instancia)

→ α (constante de normalización)

→ π - valor (cambia cuando el padre de x se instancia)

Un nodo se instancia ⟹ manda λ-mensaje a su padre y π-mensaje a sus hijos

INCISO

$P^*(+x) = \alpha \cdot 0{,}2 \implies \alpha = \dfrac{0{,}2}{0{,}2 + 0{,}4}$

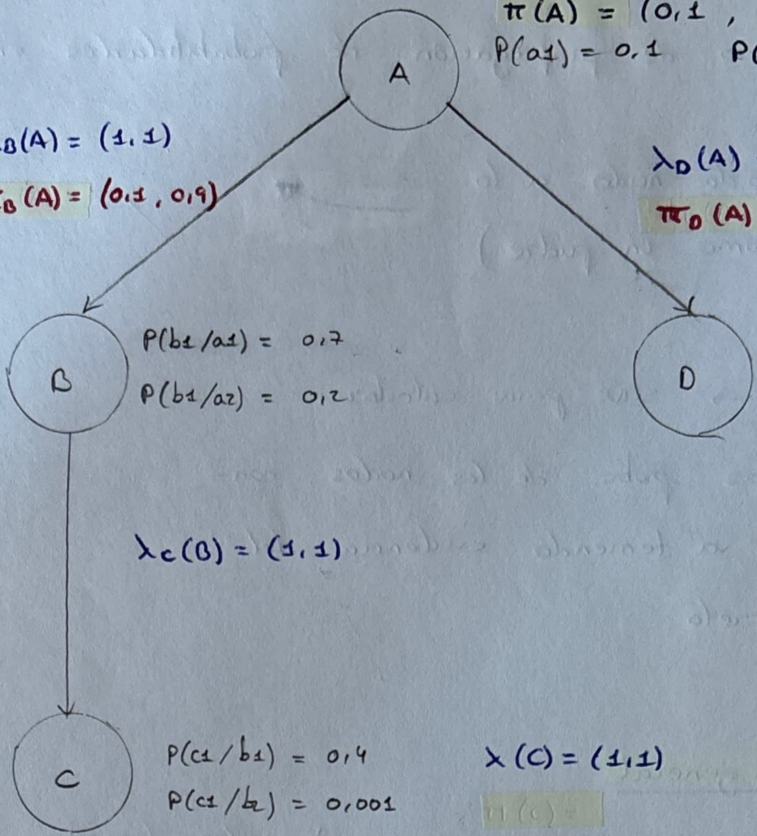
$P^*(\neg x) = \alpha \cdot 0{,}4 \implies \alpha = \dfrac{0{,}4}{0{,}2 + 0{,}4}$

¡5!

$\lambda(A) = (1, 1)$

$\pi(A) = (0{,}1\ ,\ 0{,}9)$

$P(a_1) = 0{,}1 \qquad P(A) = (0{,}1\ ,\ 0{,}9)$



$\begin{cases} 0{,}1 \cdot 1 \\ 0{,}9 \cdot 1 \end{cases}$

$\lambda_B(A) = (1, 1)$

$\pi_B(A) = (0{,}1\ ,\ 0{,}9)$

$\lambda_D(A) = (1, 1)$

$\pi_D(A) = (0{,}1\ ,\ 0{,}9)$

$\lambda(B) = (1, 1)$

$\pi(B)$

$P(b_1/a_1) = 0{,}7$

$P(b_1/a_2) = 0{,}2$

$P(d_1/a_1) = 0{,}8$

$P(d_1/a_2) = 0{,}4$

$\lambda(D) = (1, 1)$

$\pi(D)$

$\lambda_C(B) = (1, 1)$

$P(c_1/b_1) = 0{,}4$

$P(c_1/b_2) = 0{,}001$

$\lambda(C) = (1, 1)$

$\pi(C) =$

1. Probabilidades condicionadas de cada nodo dados sus padres
2. Probabilidades a priori de todos los nodos

+ Inicialización A B C
+ B ha recibido un π-mensaje de A $(\pi_B(A) = (0{,}1\ ,\ 0{,}9))$

→ Aplicamos el C

$$\pi(B) = (0{,}7 \cdot 0{,}1 + 0{,}2 \cdot 0{,}9\ ,\ 0{,}3 \cdot 0{,}1 + 0{,}8 \cdot 0{,}9) = \boxed{(0{,}25\ ,\ 0{,}75)}$$

$$P^*(B) = \begin{cases} \alpha \cdot 1 \cdot 0{,}25 \\ \alpha \cdot 1 \cdot 0{,}75 \end{cases} \Rightarrow \boxed{P^*(B) = (0{,}25\ ,\ 0{,}75)}$$

$$\pi_C(B) = \boxed{(0{,}25,\ 0{,}75)}$$

Ahora, si se instancia una variable pasaríamos a un nuevo estado $S_1, S_2, \dots$ actualizando el árbol.

Fórmulas para el cálculo de λ y π-mensajes, λ y π-valores y probabilidades P*:

1. Si B es un hijo de A, B tiene k valores posibles y A m valores posibles, entonces para $j=1,2,...,m$, el λ-mensaje de B a A viene dado por;

$$\lambda_B(a_j) = \sum_{i=1}^{k} P(b_i / a_j)\lambda(b_i).$$

2. Si B es hijo de A y A tiene m valores posibles, entonces para $j=1,2,...,m$, el π-mensaje de A a B viene dado por;

$$\pi_B(a_j) = \begin{cases} \pi(a_j) \prod\limits_{\substack{C\in s(A)\\ C\neq B}} \lambda_C(a_j) & si\ A\ no\ ha\ tomado\ ningún\ valor \\ 1 & si\ A = a_j \\ 0 & si\ A \neq a_j \end{cases}$$

donde $s(A)$ denota al conjunto de hijos de A.

3. Si B tiene k valores posibles y $s(B)$ es el conjunto de los hijos de B, entonces para $i=1,2,...,k$, el λ-valor de B viene dado por;

$$\lambda(b_i) = \begin{cases} \prod\limits_{C\in s(B)} \lambda_C(b_i) & si\ B\ no\ ha\ tomado\ ningún\ valor \\ 1 & \text{si } B = b_i \\ 0 & \text{si } B \neq b_i. \end{cases}$$

4. Si A es padre de B, B tiene k valores posibles y A tiene m valores posibles, entonces para $i=1,2,...,k$, el π-valor de B viene dado por;

$$\pi(b_i) = \sum_{j=1}^{m} P(b_i / a_j)\, \pi_B(a_j).$$

5. Si B es una variable con k posibles valores, entonces para $i = 1,2,...,k$ la probabilidad a posteriori basada en las variables instanciadas se calcula como:

$$P^*(b_i) = \alpha\, \lambda(b_i)\, \pi(b_i).$$

ALGORITMO:

1. Inicialización.

A. Inicializar todos los λ-mensajes y λ-valores a 1.
B. Si la raíz A tiene m posibles valores, entonces para $j=1,...,m$, sea
$\pi(a_j) = P(a_j)$.
C. Para todos los hijos B de la raíz A, hacer
Enviar un nuevo π-mensaje a B usando la fórmula 2.
(En ese momento comenzará un flujo de propagación debido al procedimiento de actualización C).

Cuando una variable se instancia o una variable recibe un λ o π-mensaje, se usa uno de los siguientes procedimientos de actualización;

Actualización.

A. Si una variable B se instancia a un valor b_j, entonces

BEGIN
1. Inicializar $P^*(b_j) = 1$ y $P^*(b_i) = 0$, para todo $i \neq j$.
2. Calcular $\lambda(B)$ usando la fórmula 3.
3. Enviar un nuevo λ-mensaje al padre de B usando la fórmula 1.
4. Enviar nuevos π–mensajes a los hijos de B usando la fórmula 2.
END

B. Si una variable B recibe un nuevo λ-mensaje de uno de sus hijos y la variable B no ha sido instanciada todavía, entonces,

BEGIN
1. Calcular el nuevo valor de $\lambda(B)$ usando la fórmula 3.
2. Calcular el nuevo valor de $P^*(B)$ usando la fórmula 5.
3. Enviar un nuevo λ-mensaje al padre de B usando la fórmula 1.
4. Enviar nuevos π–mensajes a los otros hijos de B usando la fórmula 2.
END.

C. Si una variable B recibe un nuevo π-mensaje de uno de sus padres y la variable B no ha sido instanciada todavía, entonces,

BEGIN
1. Calcular el nuevo valor de $\pi(B)$ usando la fórmula 4.
2. Calcular el nuevo valor de $P^*(B)$ usando la fórmula 5.
3. Enviar nuevos π–mensajes a los hijos de B usando la fórmula 2.
END.

a) Un grafo acíclico conexo y dirigido $G=(V,E)$ y una distribución de probabilidad conjunta P definida sobre V se dice que cumple la h.i.c si

$$\forall X \in V \text{ y } \forall Y \in V - \{X \cup de(X) \cup pa(X)\}$$

$$X \text{ es independiente de } Y \text{ dado } pa(X)$$

Para

d) cualquier conjunto de variables aleatorias, la probabilidad conjunta se puede calcular a partir de las probabilidades condicionadas:

$$P(x_1, \dots, x_n) = \prod_{x_i} P(x_i / pa(x_i))$$

¡Dada una red bayesiana!

Definición Formal Red Bayesiana

Una red bayesiana es

variable aleatoria que toma un conjunto exhaustivo y excluyente de valores

- Un conjunto de variables proposicionales V
- Un conjunto de relaciones binarias E definidas sobre las variables V
- Una distribución de probabilidad conjunta P sobre las variables V

tales que:

- (V, E) forman un grafo acíclico, conexo y dirigido G
- (G, P) cumplen la hipótesis de independencia condicional