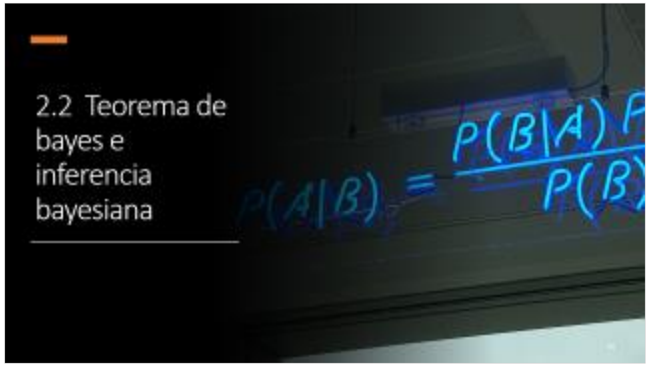
Redes Neuronales (y Bayesianas) LDS1081

Juan Manuel Ahuactzin Larios
juan.ahuactzin@udlap.mx

Todas las imágenes de este curso fueron obtenidas de Pexels, Pxhere y Microsoft Powerpoint: https://www.pexels.com/, https://pxhere.com/

Contenido: Da clic en la sección que quieres consultar.



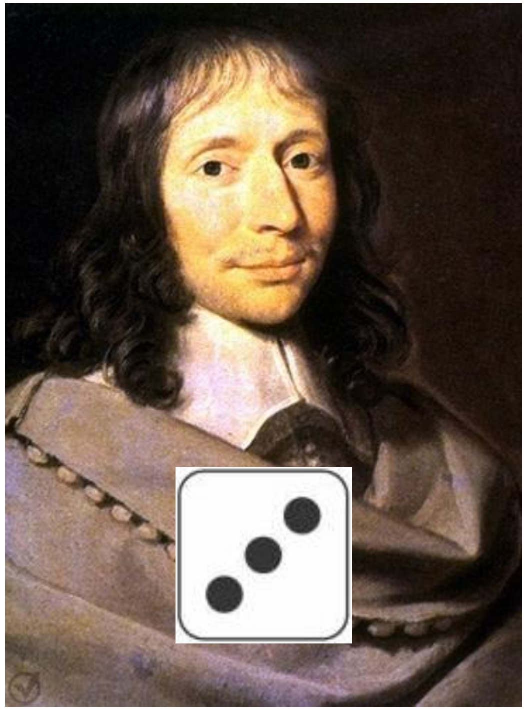
2.1 Incompetud vs Incertidumbre

Juan Manuel Ahuactzin Larios

juan.ahuactzin@udlap.mx

Experimentos aleatorios

1654



Antoine de Gombard

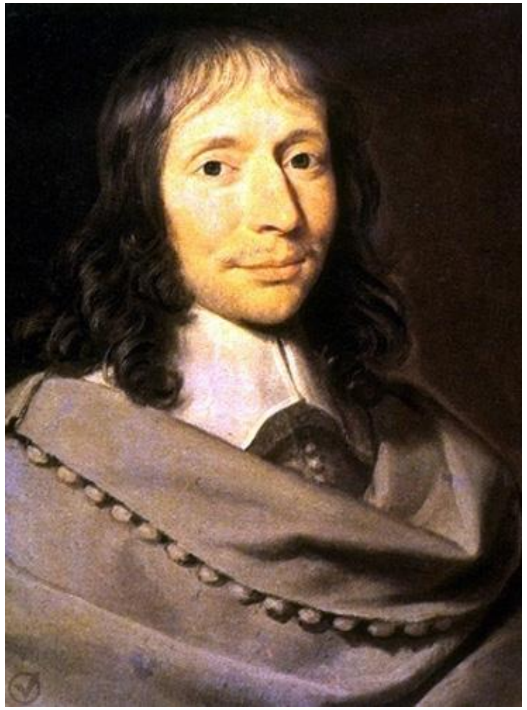
Blaise Pascal

?

Experimentos aleatorios

Experimentos aleatorios

1654



Antoine de Gombard

Blaise Pascal

Pierre de Fermat

Experimentos aleatorios

Probabilidad parte de las matemáticas que se encarga del estudio de los fenómenos aleatorios.

Experimentos aleatorios

Dos tipos de fenómenos o experimentos:

1. **Determinista**: produce el mismo resultado si lo ejecutamos bajo las mismas condiciones.
2. **Aleatorio**: no siempre produce el mismo resultado bajo las mismas condiciones.

Experimentos aleatorios

Experimentos que cumplan con dos condiciones:

1. El experimento debe poder ser repetible bajo las mismas condiciones iniciales.
2. El resultado de cualquier ensayo del experimento es variable y depende del azar o de algún mecanismo aleatorio.

Luis Rincón. Introducción a la probabilidad. Departamento de matemáticas, Facultad de Ciencias UNAM. Agosto 2014. (Visitado el 08/06/2021) http://www.economia.unam.mx/biblioteca/Pdf/Prob1Rinc%C3%B3n.pdf

Experimentos aleatorios

Clasifica los eventos

https://docs.google.com/forms/d/e/1FAIpQLSfN_3zMljHNpQEDtyna8xECvgKQsZwBjQL_xAER_5ISLs01cQ/viewform

Luis Rincón. Introducción a la probabilidad. Departamento de matemáticas, Facultad de Ciencias UNAM. Agosto 2014. (Visitado el 08/06/2021) http://www.economia.unam.mx/biblioteca/Pdf/Prob1Rinc%C3%B3n.pdf

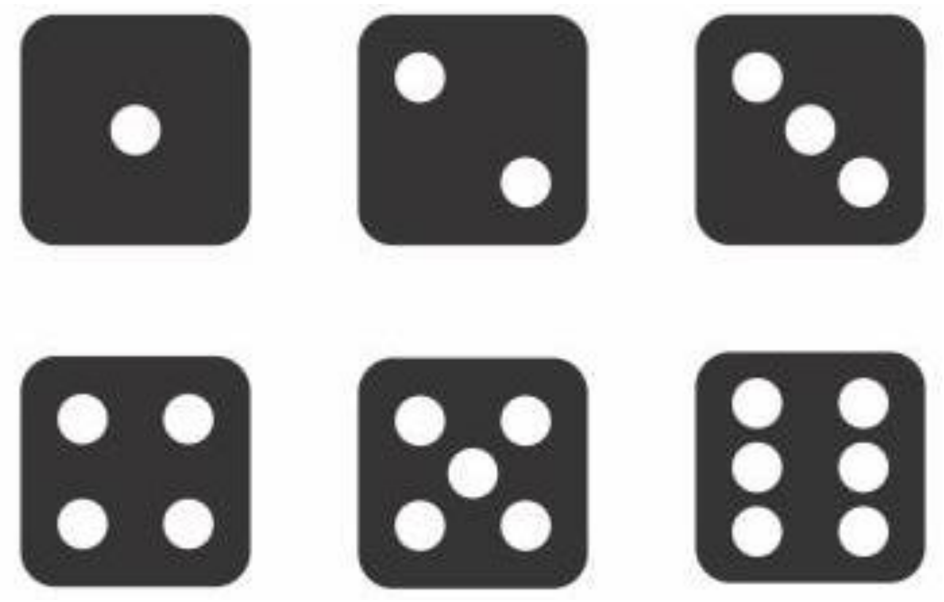
Experimentos aleatorios

Da ejemplos de eventos deterministas y aleatorios

https://docs.google.com/forms/d/e/1FAIpQLSdVDgWZ1BKYy0Ebqv4wGSaKqdlFWMozyS6PWSC0OotgO3bQhA/viewform

Espacio muestral

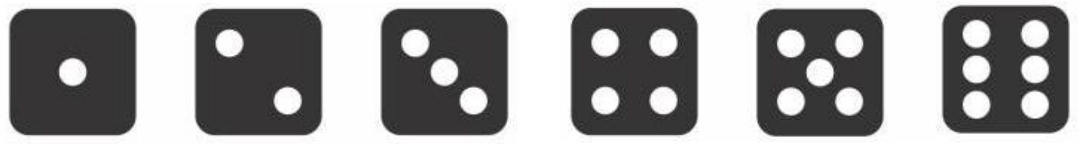
En un evento aleatorio no podemos saber el resultado pero si el conjunto de todos los resultados posibles.



Espacio muestral

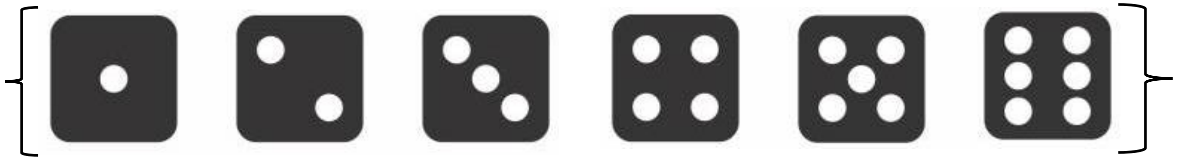
Espacio muestral es el **conjunto** de todos los resultados posibles del experimento y se denota por **Ω** (Omega mayúscula).

A **un resultado** particular del experimento se le denota por **ω** (omega minúscula).

$\Omega =$  Es un **conjunto**

$\omega =$  Es un **elemento** del conjunto

Espacio muestral

$\Omega =$ 

Es el conjunto de todos los resultados de **lanzar un dado**.

$\omega =$ 

Es el **resultado** de lanzar un dado.

Espacio muestral

Un **evento** (**o suceso**) es un **subconjunto** del espacio muestral **Ω** y se denota por **Ω** y una letra como mayúscula como subíndice por ejemplo $\boldsymbol{\Omega_A}$ o simplemente por una letra mayúscula, por ejemplo A.

$\Omega_A =$  Es un evento

¿El espacio muestral **Ω** es un evento**?**

Espacio muestral

Un **evento** está asociado a la ocurrencia de un **suceso.**

$\Omega_A = \{$  $\}$

¿Cómo podemos llamar a este evento?

R= Obtener un número par.

Espacio muestral

El espacio muestral de un experimento aleatorio no es único, depende del interés del observador.

3714
0 0 0 0 0 0
SERIE 0
Sorteo
MAYOR
Un total de
54
MILLONES
de pesos en premios*
Carla Elena
Lotería
Nacional
18
CON UN PREMIO MAYOR
MILLONES
de pesos en 3**series
MARTES 11 DE JUNIO DE 2019
Valor $25.00
DR. CARLOS MANUEL URZÚA MACÍAS
PDTE. DE LA JUNTA DIRECTIVA
LIC. ERNESTO PRIETO ORTEGA
DIRECTOR GENERAL
12799 999999119
VIGÉSIMO 18
Lea aviso importante al reverso

Espacio muestral

El espacio muestral de un experimento aleatorio no es único, depende del interés del observador.

Participan 60 mil billetes numerados del 00000 al 60000.

$$\Omega = \{00000, 00001, 00002, \ldots 60000\}$$

$$\Omega = \{0, 1, 2, \ldots 60000\}$$

$$\Omega = \{\text{ganar, perder}\}$$

Espacio muestral

Un experimento aleatorio consiste en observar el tiempo en el que hay un corte de electricidad en un fraccionamiento a partir de una fecha y hora dada.

Si se consideran mediciones continuas en el tiempo el espacio muestral es:

$$\Omega = [0, \infty)$$

Y un evento

$$\Omega_A = [1,2]$$

Espacio muestral

¿Cuál es el espacio muestral?

https://docs.google.com/forms/d/e/1FAIpQLSeKzOICSfohmaVsj9lwU1zHF3c-61vtnXSRTZnWZiHuIglkpA/viewform

Espacio muestral

Un **evento simple** consta de un solo elemento del espacio muestral.

$\Omega_A=$ {  }

Un **evento compuesto** consta de más de un elemento del espacio muestral.

$\Omega_A=$ {  }

Espacio muestral

Un **evento simple** consta de un solo elemento del espacio muestral.

$\Omega_A = \{$  $\}$

Un **evento compuesto** consta de más de un elemento del espacio muestral.

$\Omega_A = \{$  $\}$

Espacio muestral

Evento simple

$\Omega_A = \{$  $\}$ Obtener un 2

$\Omega_B = \{$  $\}$ Obtener un número mayor o igual que 6

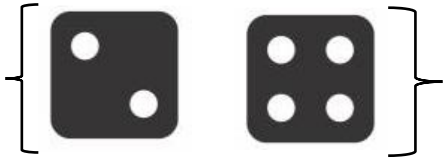
$\Omega_C = \{$  $\}$ Obtener un número divisible entre 2 y 3

Espacio muestral

Evento compuesto

$\Omega_A=$  Obtener un número par

$\Omega_B=$  Obtener un número non

$\Omega_C=$  Obtener un número par menor que 5

Espacio muestral

Evento imposible (vacío)

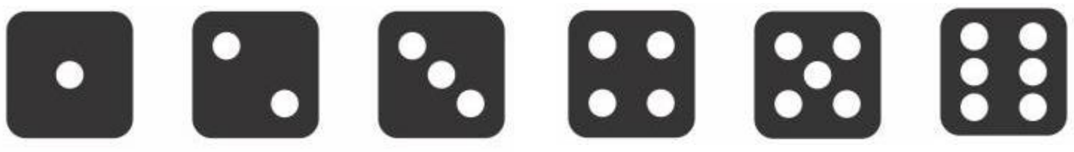
$\Omega_A = \{ \quad \}$ Obtener un número divisible entre 2 y 5

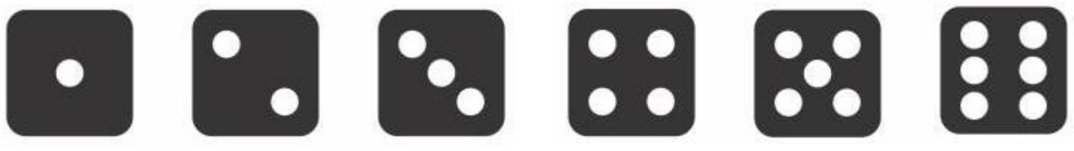
$\Omega_B = \{ \quad \}$ Obtener un número par menor que 1

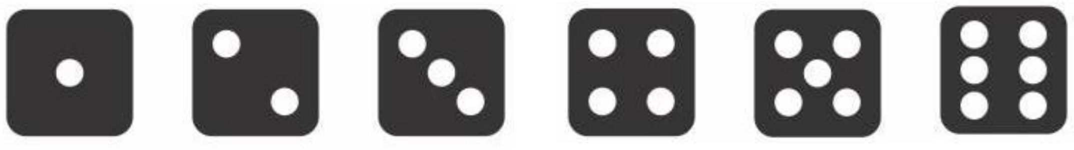
$\Omega_C = \{ \quad \}$ Obtener un número mayor a 6

Espacio muestral

Evento seguro

$\Omega_A = \{$  $\}$ Obtener un número menor que 7

$\Omega_B = \{$  $\}$ Obtener un número mayor o igual a 1

$\Omega_B = \{$  $\}$ Obtener un número entre 1 y 6

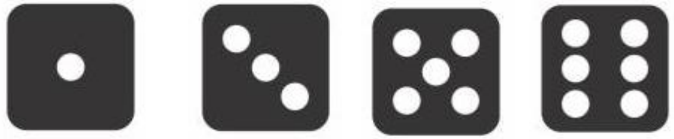
Espacio muestral

Evento contrario

$\Omega_A=$  Obtener un número par

$\Omega_A^c=$  Obtener un número non

$\Omega_B=$  Obtener un número par menor que 5

$\Omega_A^c=$  Obtener un número que no sea par y menor que 5

Espacio muestral

Tipos de evento

Clasifica los tipos de evento

https://docs.google.com/forms/d/e/1FAIpQLScIaxcMXSiO6jy96CZrnpSkIRkHAFuOQSFjlvyZWUm6Hs9yhw/viewform

Operaciones con conjuntos

Para calcular la **probabilidad** de ocurrencia de cada evento es indispensable identificar cada uno de ellos.

A partir de una colección de eventos, se pueden obtener otros más.

$\Omega_A = \{$  $\}$ Haber obtenido un número par

$\Omega_{A_1} = \{$  $\}$ Haber obtenido un número par y primo.

Operaciones con conjuntos

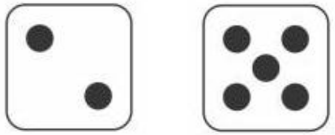
$\Omega = \{$  $\}$

Es el conjunto de todos los resultados de **lanzar un dado**.

Mi número es primo

$\Omega_A = \{$  $\}$

Mi número es blanco

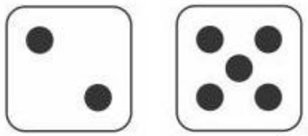
$\Omega_{A_1} = \{$  $\}$

Operaciones con conjuntos

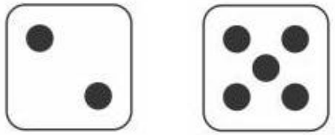
$\Omega = \{$  $\}$

Es el conjunto de todos los resultados de **lanzar un dado**.

Mi número es blanco

$\Omega_B = \{$  $\}$

Mi número es primo

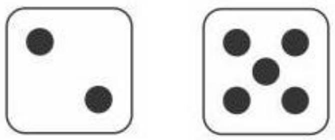
$\Omega_{B_1} = \{$  $\}$

Operaciones con conjuntos

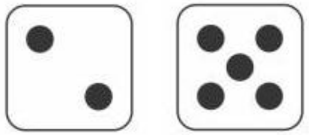
Sabiendo que mi número es primo
que proporción es blanco.

$\Omega_A = \{$  $\}$ 3

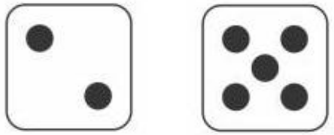
$$\frac{2}{3}$$

$\Omega_{A_1} = \{$  $\}$ 2

Sabiendo que mi número blanco
que proporción es primo.

$\Omega_B = \{$  $\}$ 2

$$\frac{1}{1}$$

$\Omega_{B_1} = \{$  $\}$ 2

Operaciones con conjuntos

Conjunto potencia

El conjunto potencia de Ω el conjunto constituido por todos lo subconjuntos de Ω y se denota por $pot(\Omega)$

Ejemplo:

Ω ={Juan, Miriam, Roberto}

2^{Ω} = {∅, {Juan}, {Miriam}, {Roberto}, {Juan, Miriam}, {Juan, Roberto}, {Miriam, Roberto}, Ω}

La cardinalidad de $pot(\Omega)$es $2^{\#\Omega}$

Operaciones con conjuntos

Producto cartesiano

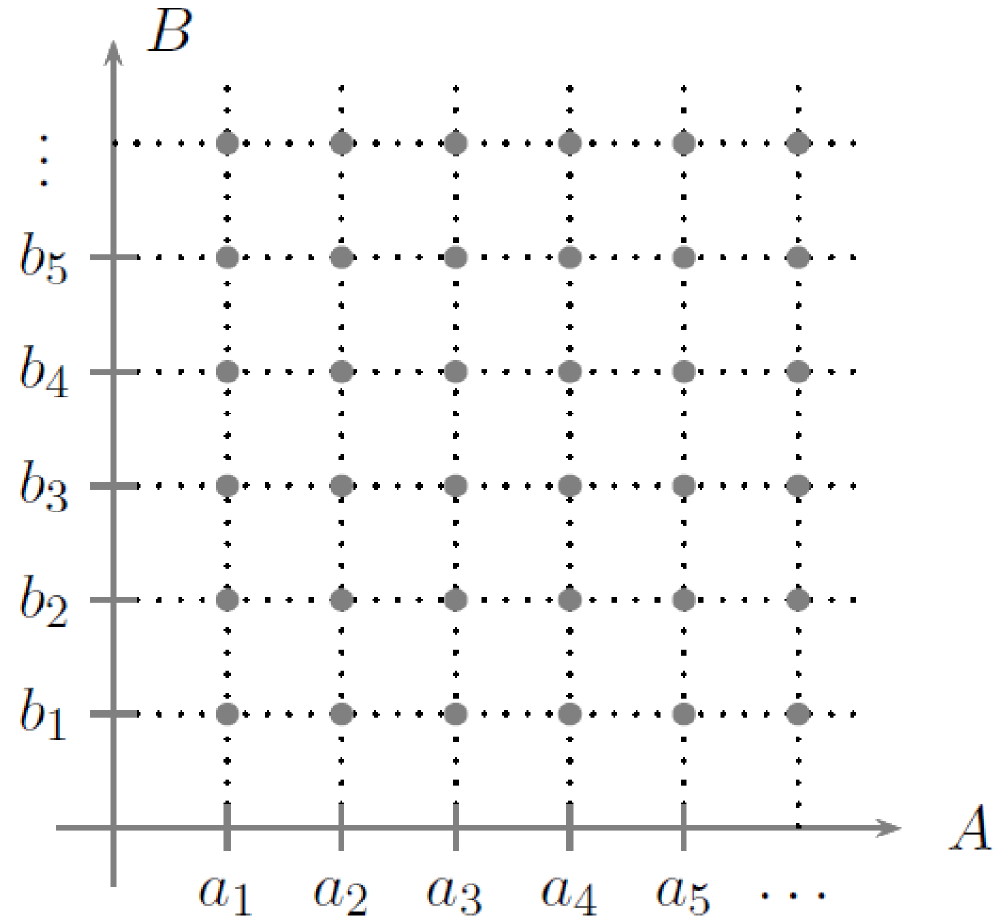
El producto cartesiano de dos conjuntos A y B se define como el conjunto de todas las parejas ordenadas de (a,b) con a perteneciente a A y b perteneciente a B y se denota por " $\times$ "

$$A \times B = \{(a, b) : a \in A \text{ y } b \in B\}.$$

La cardinalidad de $\mathrm{A} \times B$ es $Card(a) * Card(b)$

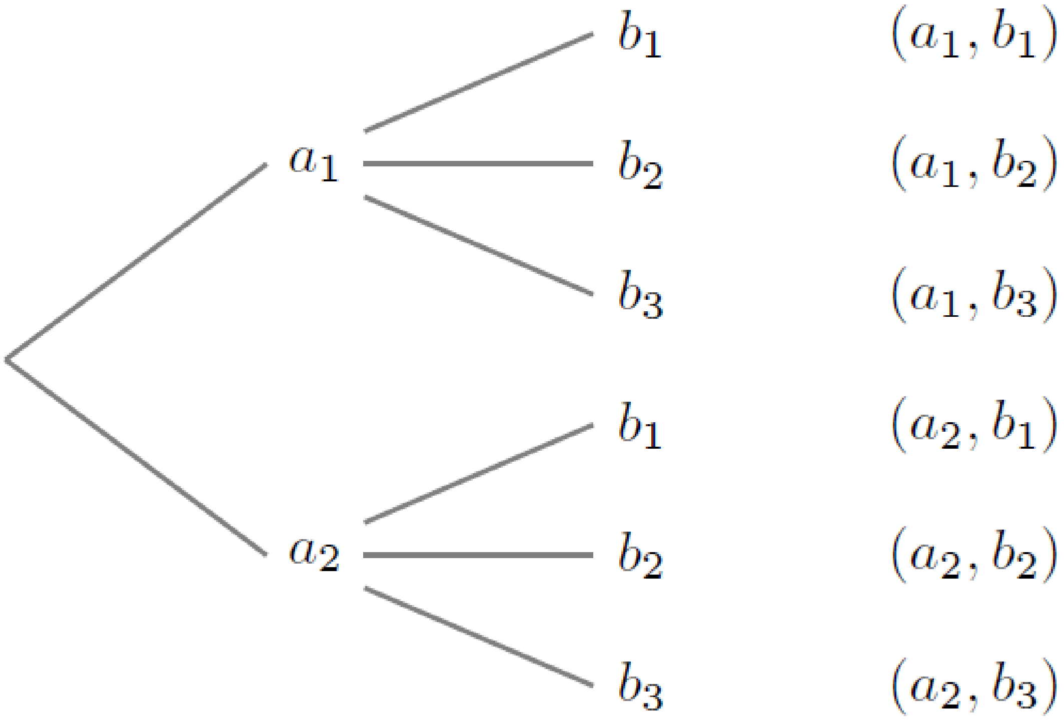
Operaciones con conjuntos

Representación del Producto cartesiano



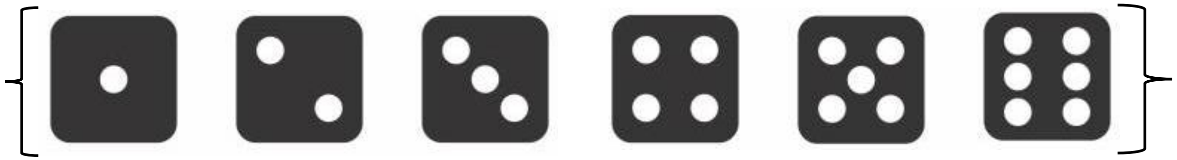
Operaciones con conjuntos

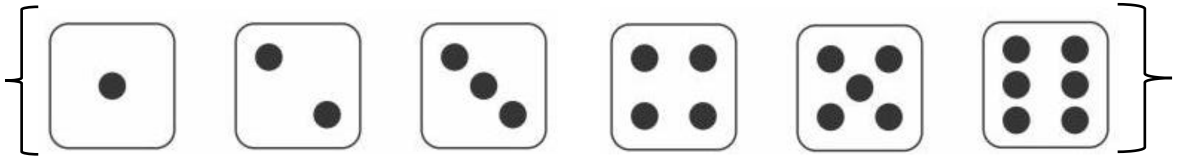
Representación del Producto cartesiano



Operaciones con conjuntos

Representación del Producto cartesiano

$\Omega_N =$ 

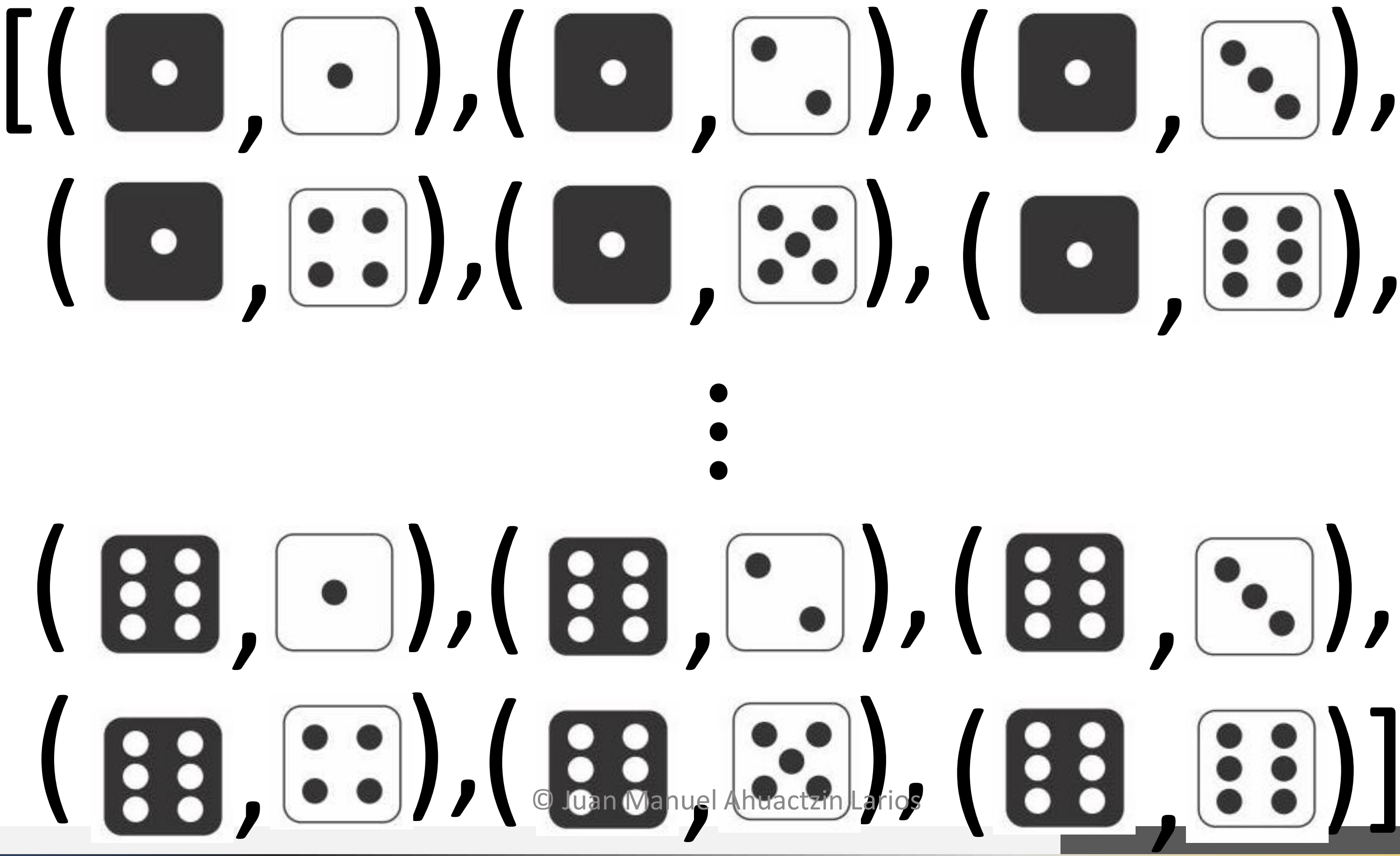
$\Omega_B =$ 

$$\Omega_N \times \Omega_B$$

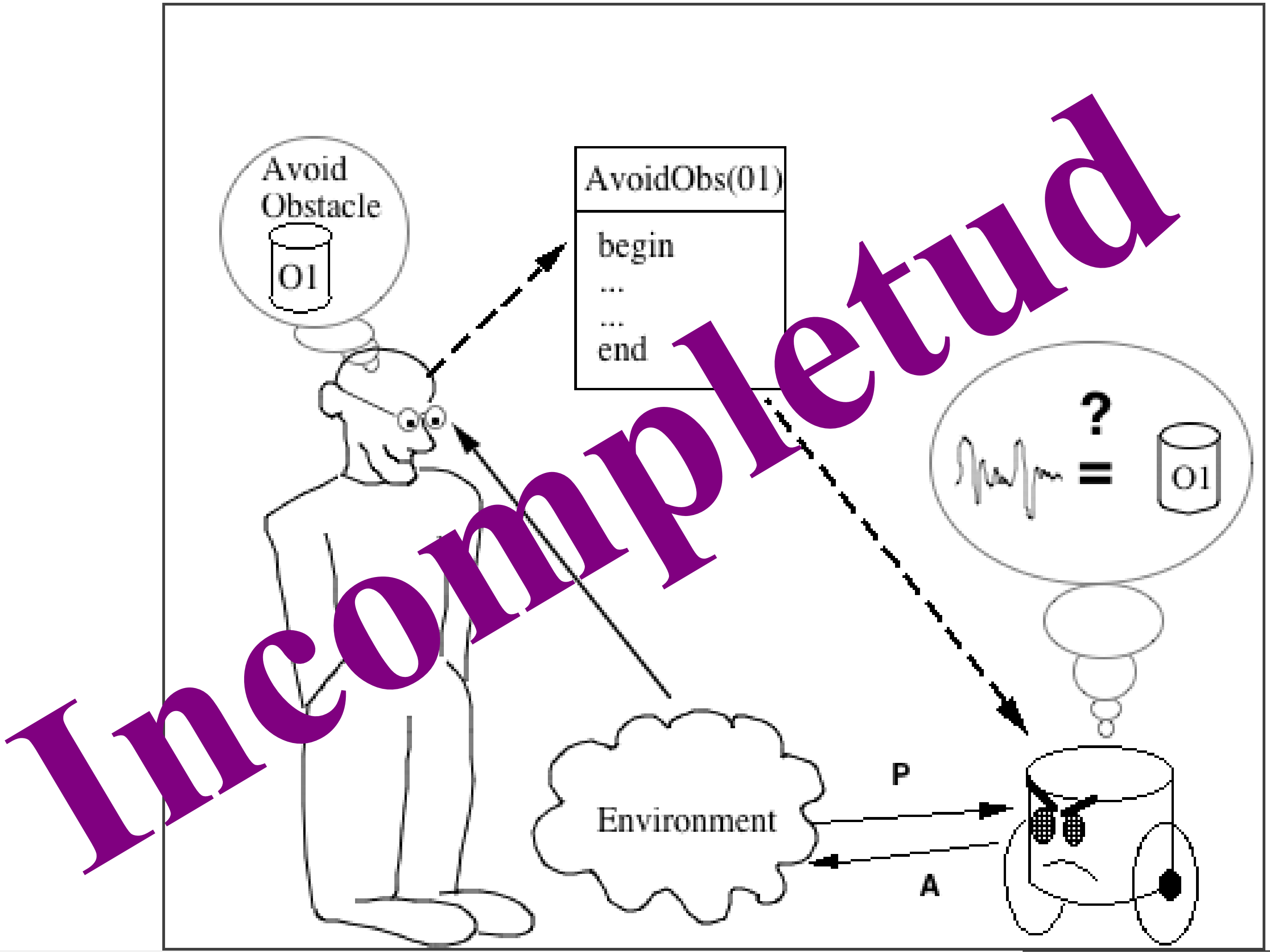
Operaciones con conjuntos

Representación del Producto cartesiano

$\Omega_N \times \Omega_B =$



Paradigma lógico



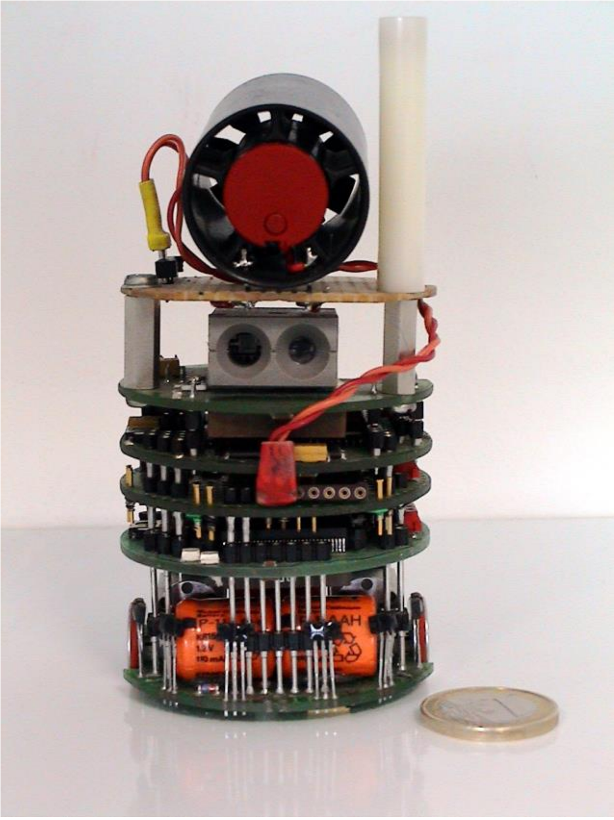
	Izquierda	Derecho	Derecha
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0

Photo Alain Herzog

Aprendiendo comportamientos reactivos

Lebeltel, O., Bessière, P., Diard, J. & Mazer, E. (2004) Bayesian Robot Programming; *Autonomous Robots*, Vol. 16, p. 49-79

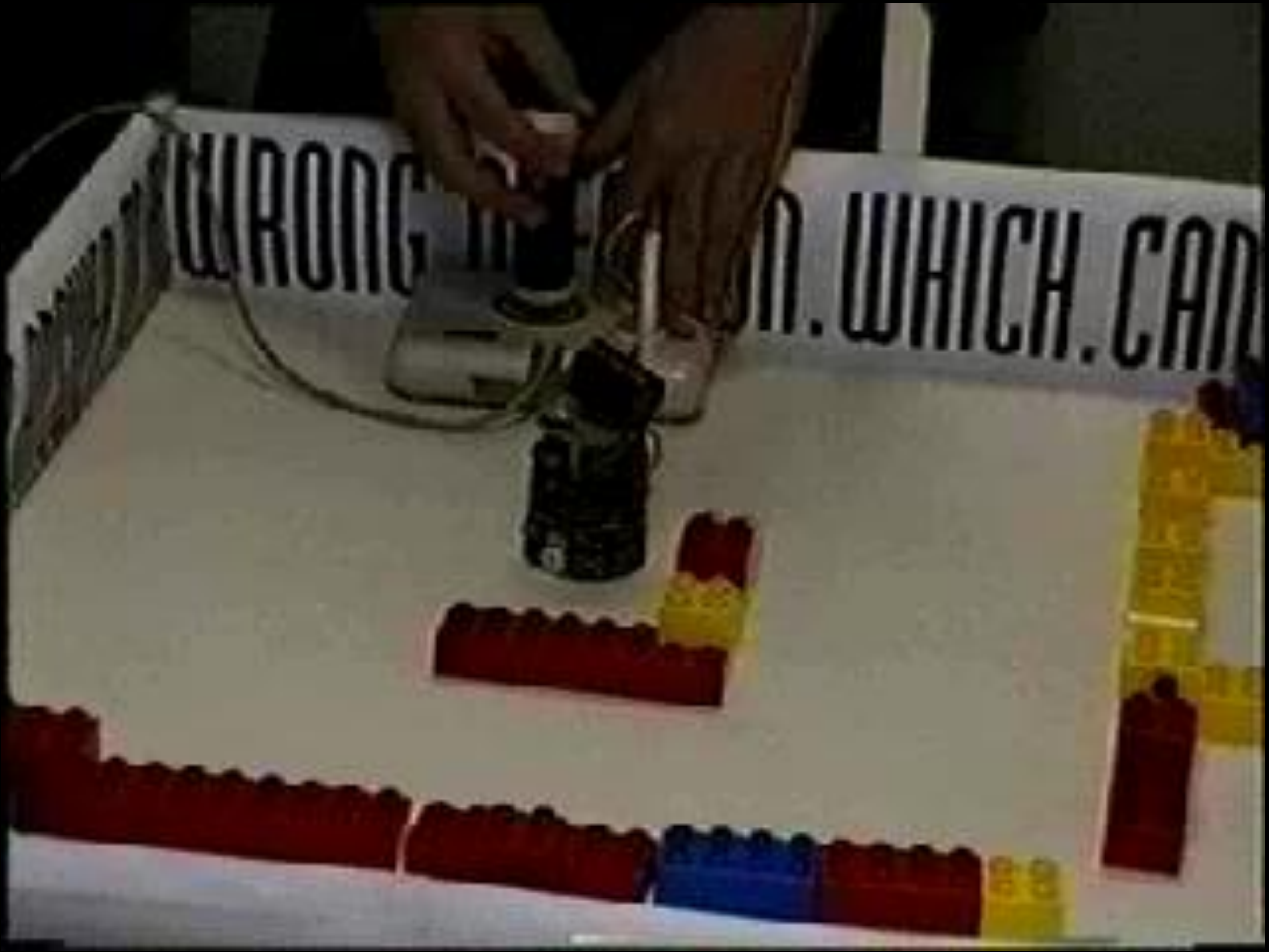
Lebeltel, O. (1999) *Programmation bayésienne des robots*; Thèse INPG

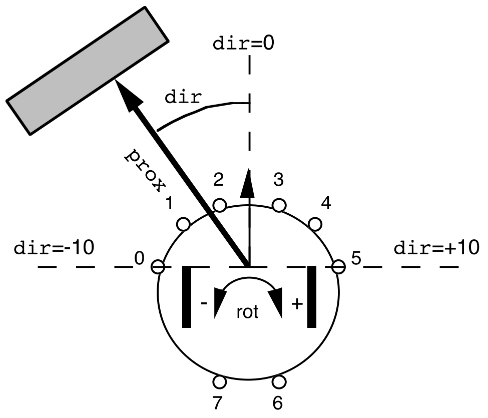


Khepera

- Avoiding Obstacle
- Contour Following
- Piano mover
- Phototaxy
- etc.


WRONG,
WHICH.


WRONG
WHICH



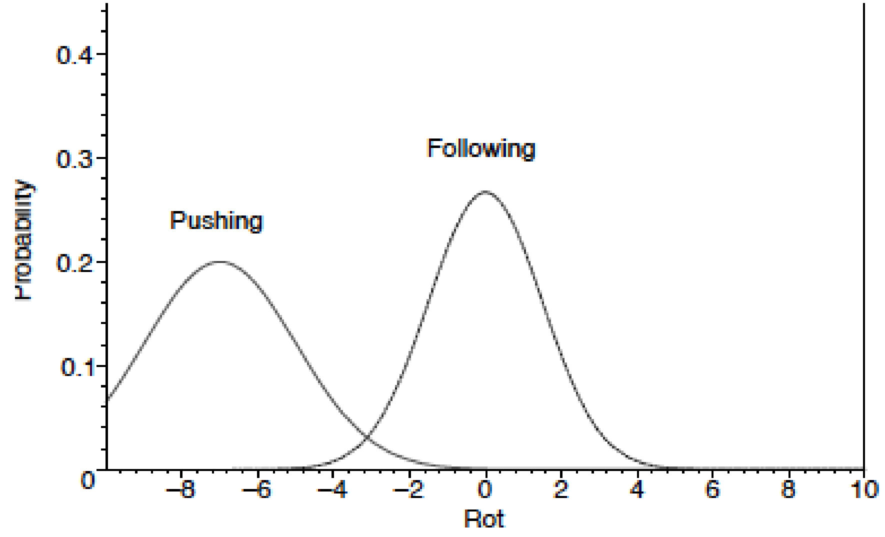


Figure 4.4: $\mathbf{P}(Rot \mid [Dir = -10] \wedge [Prox = 13])$

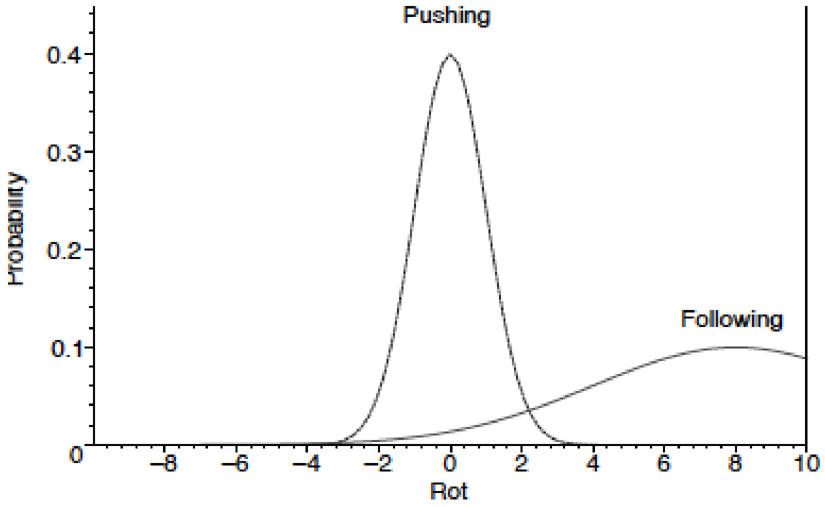


Figure 4.5: $\mathbf{P}(Rot \mid [Dir = 0] \wedge [Prox = 15])$

Principios

Incompletud

Aprendizaje bayesiano

Conocimientos previos
+
Datos experimentales
=
Representación probabilística

Incertidumbre

Inferencia bayesiana

$$P(a)+P(\neg a)=1$$

$$P(a\wedge b)=P(a)P(b\mid a)=P(b)P(a\mid b)$$

Decisión

Silogismos

a : “x es divisible por 9”
b: “x es divisible por 3”

- Silogismos lógicos :
  - Modus Ponens: $a \wedge [a \Rightarrow b] \mapsto b$
  - Modus Tollens: $\neg b \wedge [a \Rightarrow b] \mapsto \neg a$
- Silogismos probabilísticos :
  - Modus Ponens: $\mathrm{P}(b \mid a) = 1$

$$\mathrm{P}(b \mid a) = 1 \Leftrightarrow \mathrm{P}(\neg a \mid \neg b) = 1$$

  - Modus Tollens:

$$\mathrm{P}(b \mid a) = 1 \Rightarrow \mathrm{P}(a \mid b) \geq \mathrm{P}(a)$$

$$\mathrm{P}(b \mid a) = 1 \Rightarrow \mathrm{P}(b \mid \neg a) \leq \mathrm{P}(b)$$

$$P(b|a)=1 \Rightarrow P(a|b) \geq P(a)$$

La probabilidad de que sea divisible entre 9 sabiendo que es divisible entre 3 es mayor que la probabilidad de que sea divisible entre 9.

a : “x es divisible por 9”
b: “x es divisible por 3”

No sé nada:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, …

Entonces la probabilidad de que x sea divisible por 9, $P(a) = \frac{3}{30} = \frac{1}{10}$

1, 2, 3, 4, 5, 6, 7, 8, **9**, 10, 11, 12, 13, 14, 15, 16, 17, **18**, 19, 20, 21, 22, 23, 24, 25, 26, **27**, 28, 29, 30, … } 3 de 30 números

Sé que x es divisible por 3:
3, 6, 9, 12, 15, 18, 21, 24, 27, 30, … } 10 números

Entonces la probabilidad de que x sea divisible por 9 sabiendo que es divisible por 3, $P(a\,|b) = \frac{3}{10}$
por que **3**, **6**, **9**, **12**, **15**, **18**, **21**, **24**, **27**, **30**, … } 3 de los 10 números

Cláramente: $P(b|a)=1 \Rightarrow P(a|b) \geq P(a) \;=\; \frac{3}{10} \geq \frac{1}{10}$

a : “x es divisible por 9”
b: “x es divisible por 3”

P(b|a)=1 Þ P(b|Øa) £ P(b)

La probabilidad de que sea divisible entre 3 sabiendo que no es divisible entre 9 es menor que la probabilidad de que sea divisible entre 3.

Si no sé nada:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, …

Entonces la probabilidad de que x sea divisible por 3, $P(b) = \frac{10}{30} = \frac{1}{3}$

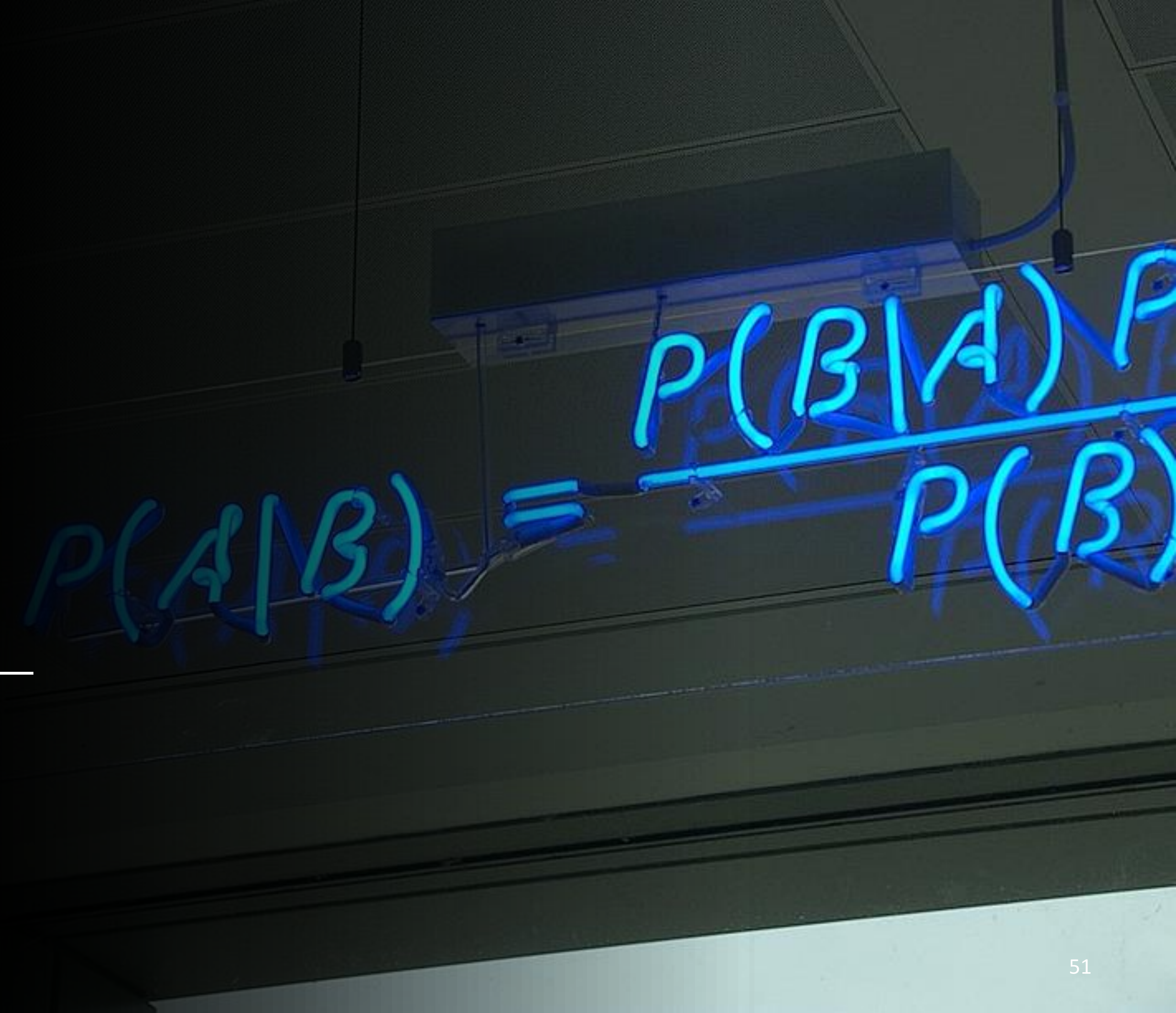
1, 2, **3**, 4, 5, **6**, 7, 8, **9**, 10, 11, **12**, 13, 14, **15**, 16, 17, **18**, 19, 20, **21**, 22, 23, **24**, 25, 26, **27**, 28, 29, **30**, …

Sé que x no es divisible por 9 (notar que quitamos al 9, 18, 27):
1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17, 19, 20, 21, 22, 23, 24, 25, 26, 28, 29, 30, …

Entonces la probabilidad de que x sea divisible por 3 sabiendo que no es divisible por 9, $P(b \mid \neg a) = \frac{7}{27}$
because 1, 2, **3**, 4, 5, **6**, 7, 8, 10, 11, **12**, 13, 14, **15**, 16, 17, 19, 20, **21**, 22, 23, **24**, 25, 26, 28, 29, **30**, …

Cláramente: P(b|a)=1 Þ P(b|Øa) £ P(b) $= \frac{7}{27} \leq \frac{1}{3} = 0.259 \leq 0.33333$

2.2 Teorema de bayes e inferencia bayesiana



Teorema de Bayes

$$P(A|B) = \frac{P(A)\,P(B\,|A)}{P(B)} = \frac{P(A\,B)}{P(B)}$$

Teorema de Bayes

X = vive en Grenoble

Y = sabe quien es Bayard

$$P(X = f \quad Y = f) = 0.92$$

$$P(X = f \quad Y = v) = 0.01$$

$$P(X = v \quad Y = f) = 0.04$$

$$P(X = v \quad Y = v) = 0.03$$

Teorema de Bayes

X = vive en Grenoble

Y = sabe quien es Bayard

$$P(X = f \quad Y = f) = 0.92$$

$$P(X = f \quad Y = v) = 0.01$$

$$P(X = v \quad Y = f) = 0.04$$

$$P(X = v \quad Y = v) = 0.03$$

Teorema de Bayes

$$P(XY)$$

X \ Y	**0**	**1**
0	0.92	0.01
1	0.04	?

Teorema de Bayes

$$P(XY)$$

X \ Y	**0**	**1**
0	0.92	0.01
1	0.04	0.03

$$\sum_X \sum_Y P(XY) = 1.0$$

Teorema de Bayes

$$P(XY)$$
$$P(X)$$
$$P(Y)$$
$$P(X|Y)$$
$$P(Y|X)$$

X \ Y	0	1	
0	0.92	0.01	0.93
1	0.04	0.03	0.07
	0.96	0.04	

Teorema de Bayes

$$P(X)P(Y \mid X) = P(Y)P(X \mid Y)$$
$$= P(XY)$$

$$P(X \mid Y) = \frac{P(XY)}{P(Y)} = \frac{P(XY)}{\sum_X P(XY)}$$

$$P(X \mid Y) = \frac{P(X)P(Y \mid X)}{P(Y)} = \frac{P(X)P(Y \mid X)}{\sum_X P(X)P(Y \mid X)}$$

Ejercicio 04: Teorema de Bayes Bayard

Comprobar el teorema de Bayes (subir el archivo)

El teorema de Bayes para 3 variables

Ejercicio 05: Teorema de Bayes

Identificar las descomposiciones válidas

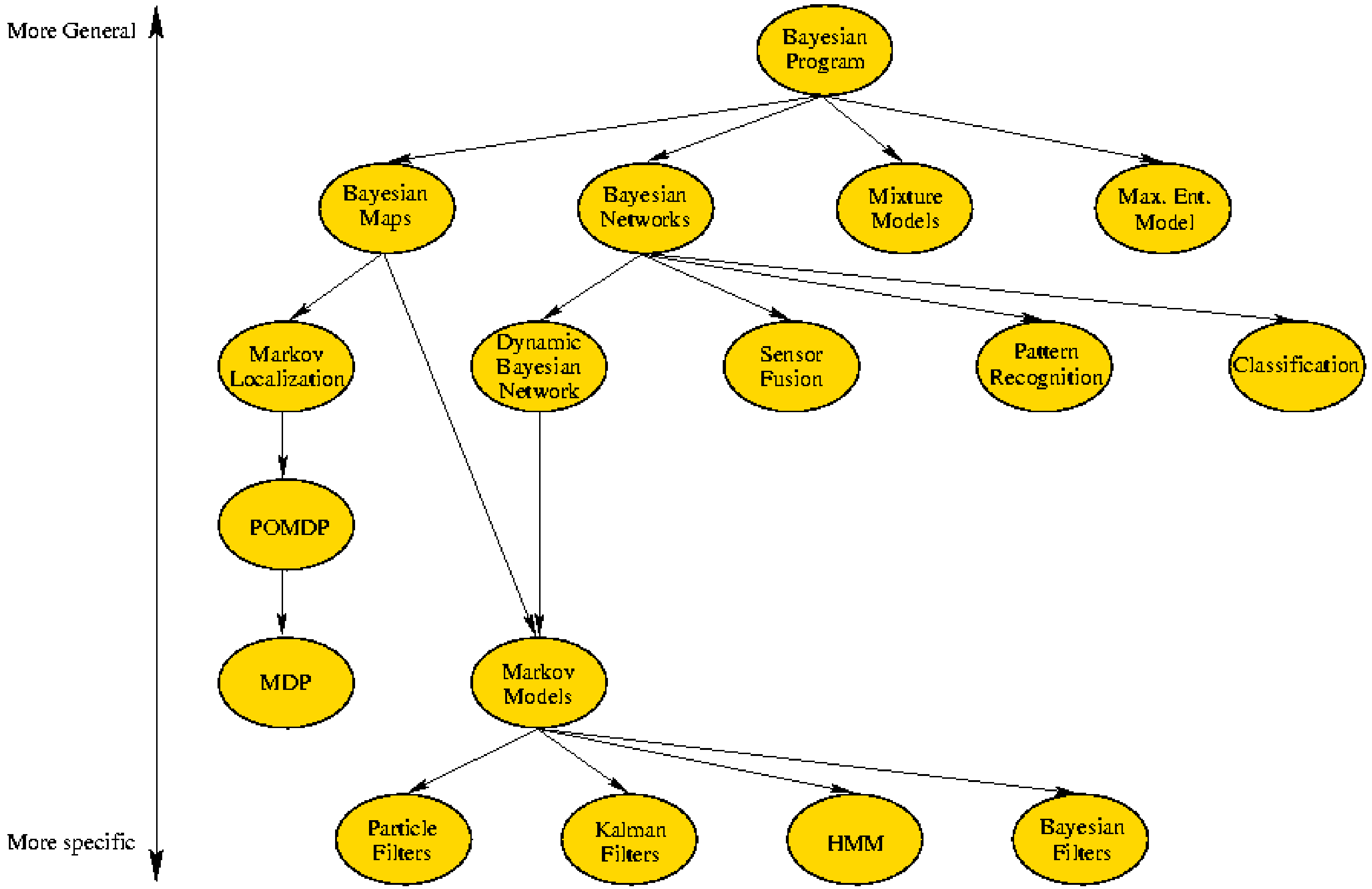
https://docs.google.com/forms/d/e/1FAIpQLSdUcYBFL6qTleZAjJqqsPeYYmRZB8a1Om828ZtbEMuVbrAVYg/viewform?usp=sf_link

Redes Neuronales (y Bayesianas) LDS1081

Juan Manuel Ahuactzin Larios
juan.ahuactzin@udlap.mx

INICIA GRABACIÓN DE VIDEO

Jerarquía de los programas bayesianos



Ejemplo 14: Inferencia con el Teorema de Bayes

2.3 Redes bayesianas

Ejercicio 06: Dibujar las redes bayesianas

https://docs.google.com/forms/d/e/1FAIpQLScCyvrMpiW4D6UetBPg34YjnhM9_OQmPp2wtpiG36RncqMbuw/viewform?usp=sf_link

Bayesian program = Description + Question

- Program
  - Description
    - *Specification*
      - Variables
      - Decomposition
      - Parametrical Forms
    - *Identification*
      - Experimental data
  - Question

$$
BiasedDice() = \left\{
\begin{array}{l}
\text{Description} \left\{
\begin{array}{l}
\text{Specification} \left\{
\begin{array}{l}
\text{Relevant Variables:} \\
Die, Points \\
\text{Decomposition:} \\
P(Die\ Points \mid \pi) = \\
\quad P(Die \mid \pi) P(Points \mid Die\ \pi) \\
\text{Parametric Forms:} \\
P(Die \mid \pi) = \begin{array}{|c|c|} \hline \mathbf{0} & \mathbf{1} \\ \hline 0.36 & 0.64 \\ \hline \end{array} \\
\\
P(Points \mid Die\ \pi) = \begin{array}{|c|c|c|c|c|c|c|} \hline Die & \multicolumn{6}{c|}{Points} \\ \cline{2-7} & \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} & \mathbf{5} & \mathbf{6} \\ \hline \mathbf{0} & 0.3 & 0.2 & 0.1 & 0.1 & 0.2 & 0.1 \\ \hline \mathbf{1} & 0.16 & 0.16 & 0.16 & 0.16 & 0.16 & 0.16 \\ \hline \end{array}
\end{array}
\right. \\
\text{Identification:} \\
\quad \text{All tables provided by the user}
\end{array}
\right. \\
\text{Question:} \\
P(Die|Points)
\end{array}
\right.
$$

Detección bayesiana de *spam*

- Clasificar los textos en 2 categorías "spam" o "no spam"
- Única información disponible: un conjunto de palabras
- Se adapta al usuario y aprende de la experiencia

Variables

$Spam$

$W_0, W_1, \ldots, \mathrm{W}_{\mathrm{N\text{-}1}}$

Probabilidad

$$P([Spam = false]) = 0.25$$

$$P([Spam = true]) = 0.75$$

Postulado de normalización

$$P([Spam = false]) + P([Spam = true]) = 1.0$$

$$\sum_{x \in X} P([X = x]) = 1.0$$

$$\sum_{X} P(X) = 1.0$$

Probabilidad condicional

$$P([W_n = false] \mid [Spam = true]) = 0.9996$$

$$P([W_n = true] \mid [Spam = true]) = 0.0004$$

$$\sum_{X} P(X \mid Y) = 1.0$$

Conjunción de variables

$$P\left(Spam \wedge W_n\right)$$

$$P\left(Spam \wedge W_0 \wedge \ldots \wedge W_n \wedge \ldots \wedge W_{N-1}\right)$$

Postulado de la conjunción

$$P(X \wedge Y) = P(X) \times P(Y \mid X)$$
$$= P(Y) \times P(X \mid Y)$$

$$P(X \mid Y) = \frac{P(X) \times P(Y \mid X)}{P(Y)}$$

$$P([Spam = true] \wedge [W_n = true]) = P([Spam = true]) \times P([W_n = true] \mid [Spam = true])$$
$$= 0.75 \times 0.0004$$
$$= 0.0003$$
$$= P([W_n = true]) \times P([Spam = true] \mid [W_n = true])$$

Regla de marginación

$$\sum_{X} P(X \wedge Y) = P(Y)$$

Distribución conjunta y preguntas (1)

$$P(Y) = \sum_{X} P(X \wedge Y)$$

$$P(X) = \sum_{Y} P(X \wedge Y)$$

$$P(Y \mid X) = \frac{P(X \wedge Y)}{\sum_{Y} P(X \wedge Y)}$$

$$P(X \mid Y) = \frac{P(X \wedge Y)}{\sum_{X} P(X \wedge Y)}$$

Distribución conjunta y preguntas (2)

$$P(Spam \wedge W_0 \wedge \ldots \wedge W_n \wedge \ldots \wedge W_{N-1})$$

$$P(Spam) = \sum_{W_0 \wedge \ldots \wedge W_{N-1}} P(Spam \wedge W_0 \wedge \ldots \wedge W_n \wedge \ldots \wedge W_{N-1})$$

$$P(W_n) = \sum_{Spam \wedge W_{i \neq n}} P(Spam \wedge W_0 \wedge \ldots \wedge W_n \wedge \ldots \wedge W_{N-1})$$

Distribución conjunta y preguntas (3)

$$P(W_n \mid [Spam = true]) = \frac{\sum_{W_{i \neq n}} P(Spam \wedge W_0 \wedge \ldots \wedge W_n \wedge \ldots \wedge W_{N-1})}{\sum_{W_0 \wedge \ldots \wedge W_{N-1}} P(Spam \wedge W_0 \wedge \ldots \wedge W_n \wedge \ldots \wedge W_{N-1})}$$

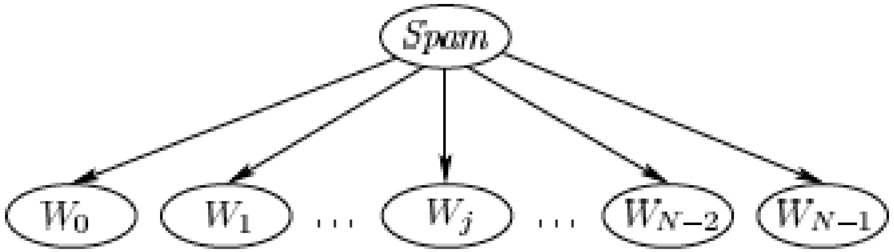
$$P(Spam \mid W_0 \wedge \ldots \wedge W_{N-1}) = \frac{P(Spam \wedge W_0 \wedge \ldots \wedge W_n \wedge \ldots \wedge W_{N-1})}{\sum_{Spam} P(Spam \wedge W_0 \wedge \ldots \wedge W_n \wedge \ldots \wedge W_{N-1})}$$

Decompocición

$$P(Spam \wedge W_0 \wedge \ldots \wedge W_{N-1})$$
$$= P(Spam) \times P(W_0 \mid Spam) \times P(W_1 \mid W_0 \wedge Spam)$$
$$\times \ldots \times P(W_{N-1} \mid W_{N-2} \wedge \ldots \wedge W_0 \wedge Spam)$$

$$P(Spam \wedge W_0 \wedge \ldots \wedge W_{N-1})$$
$$= P(Spam) \times \prod_{n=0}^{N-1} P(W_n \mid Spam)$$

Red bayesiana



Formas paramétricas (1)

$$P(Spam \wedge W_0 \wedge \ldots \wedge W_{N-1})$$

$$= P(Spam) \times \prod_{n=0}^{N-1} P(W_n \mid Spam)$$

$$P([Spam = false]) = 0.25$$

$$P(W_n \mid [Spam = false]) = \frac{Nbre_f^n}{Nbre_f}$$

$$P(W_n \mid [Spam = true]) = \frac{Nbre_t^n}{Nbre_t}$$

Formas paramétricas (2)

$$P\left(Spam \wedge W_0 \wedge \ldots \wedge W_{N-1}\right)$$

$$= P(Spam) \times \prod_{n=0}^{N-1} P(W_n \mid Spam)$$

$$P([Spam = false]) = 0.25$$

$$P(W_n \mid [Spam = false]) = \frac{1 + Nbre_f^n}{2 + Nbre_f}$$

$$P(W_n \mid [Spam = true]) = \frac{1 + Nbre_t^n}{2 + Nbre_t}$$

Identificación

$$P\left(W_n \mid [Spam = false]\right) = \frac{1 + Nbre_f^n}{2 + Nbre_f}$$

$$P\left(W_n \mid [Spam = true]\right) = \frac{1 + Nbre_t^n}{2 + Nbre_t}$$

Specification = Variables + Decomposition + Parametric Forms

- ***Variables**: la elección de las variables relevantes para el problema*
- ***Descomposición**: la expresión de la distribución de probabilidad conjunta como el producto de una distribución más simple*
- ***Formas paramétricas**: la elección de las funciones matemáticas de cada una de estas distribuciones*

Description = Specification + Identification

$$P\left(W_n \mid [Spam = false]\right) = \frac{1 + Nbre_f^n}{2 + Nbre_f}$$

$$P\left(W_n \mid [Spam = true]\right) = \frac{1 + Nbre_t^n}{2 + Nbre_t}$$

Preguntas (1)

$$P(Spam \wedge W_0 \wedge \ldots \wedge W_n \wedge \ldots \wedge W_{N-1}) = P(Spam) \times \prod_{n=0}^{N-1} P(W_n \mid Spam)$$

$$P(Spam) = \sum_{W_0 \wedge \ldots \wedge W_{N-1}} P(Spam) \times \prod_{n=0}^{N-1} P(W_n \mid Spam)$$

$$= P(Spam)$$

Preguntas (2)

$$P(W_n) = \sum_{Spam \wedge W_{i \neq n}} P(Spam) \times \prod_{n=0}^{N-1} P(W_n \mid Spam)$$

$$= \sum_{Spam} P(Spam) \times P(W_n \mid Spam)$$

$$P([W_n = true]) = \left(0.25 \times \frac{1 + Nbre_f^n}{2 + Nbre_f}\right) + \left(0.75 \times \frac{1 + Nbre_t^n}{2 + Nbre_t}\right)$$

Preguntas (3)

$$P\left(W_n \mid [Spam = true]\right) = \frac{1 + Nbre_t^n}{2 + Nbre_t}$$

$$P\left(Spam \mid [W_n = true]\right) = \frac{P(Spam) \times P\left([W_n = true] \middle| Spam\right)}{\sum_{Spam} P(Spam) \times P\left([W_n = true] \middle| Spam\right)}$$

Preguntas (4)

$$P(Spam \mid W_0 \wedge \ldots W_{N-1}) = \frac{P(Spam) \times \prod_{n=0}^{N-1} P(W_n \mid Spam)}{\sum_{Spam} P(Spam) \times \prod_{n=0}^{N-1} P(W_n \mid Spam)}$$

$$\frac{P([Spam = true] \mid W_0 \wedge \ldots \wedge W_{N-1})}{P([Spam = false] \mid W_0 \wedge \ldots \wedge W_{N-1})} = \frac{P([Spam = true])}{P([Spam = false])} \times \prod_{n=0}^{N-1} \frac{P(W_n \mid [Spam = true])}{P(W_n \mid [Spam = false])}$$

Bayesian Program = Description + Question

Program

Description

Specification

- Variables
- Decomposition
- Parametrical Forms or Recursive Question

Preliminary Knowledge π

Identification

Experimental Data δ

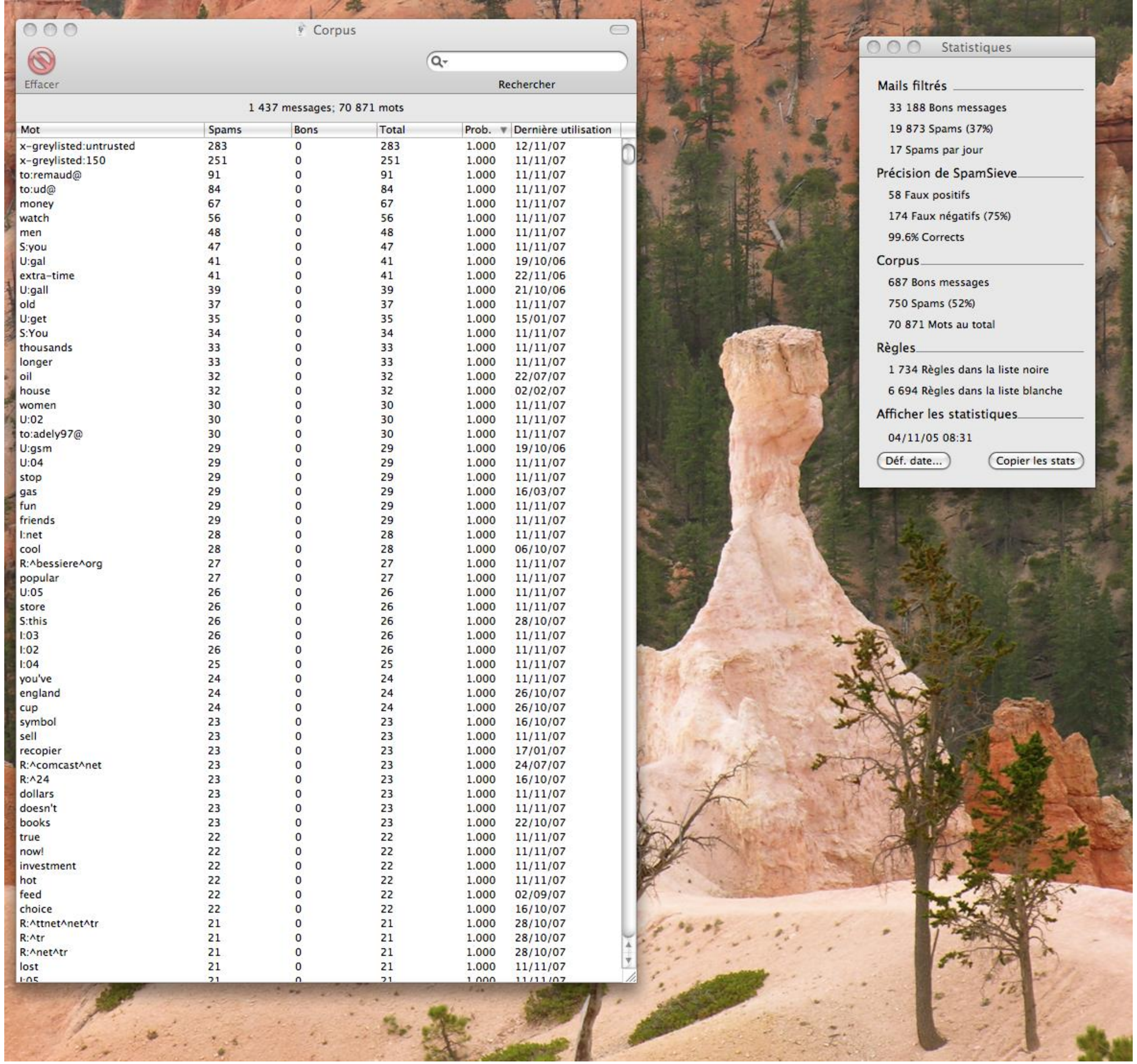
Question

Use

Bayesian Program = Description + Question

$$
\mathbf{Pr}\begin{cases}
\mathbf{Ds}\begin{cases}
\mathbf{Sp}\begin{cases}
\mathbf{Va} \rightarrow \mathit{Spam}, W_0, W_2, \ldots, W_{N-1} \\
\mathbf{Dc} \rightarrow \begin{cases}
\mathbf{P}(\mathit{Spam} \wedge W_0 \wedge \ldots \wedge W_j \wedge \ldots \wedge W_{N-1}) \\
= \mathbf{P}(\mathit{Spam}) \times \prod_{i=1}^{N} \mathbf{P}(W_i \mid \mathit{Spam})
\end{cases} \\
\mathbf{Fo} \rightarrow \begin{cases}
\mathbf{P}(\mathit{Spam}) \rightarrow \begin{cases}
\mathbf{P}([\mathit{Spam} = \text{false}]) = 0{,}25 \\
\mathbf{P}([\mathit{Spam} = \text{true}]) = 0{,}75
\end{cases} \\
\mathbf{P}(W_i \mid \mathit{Spam}) \rightarrow \begin{cases}
\mathbf{P}([W_i = \text{false}] \mid [\mathit{Spam} = \text{false}]) = 1 - \dfrac{1 + n_f^i}{2 + n_f} \\
\mathbf{P}([W_i = \text{true}] \mid [\mathit{Spam} = \text{false}]) = \dfrac{1 + n_f^i}{2 + n_f} \\
\mathbf{P}([W_i = \text{false}] \mid [\mathit{Spam} = \text{true}]) = 1 - \dfrac{1 + n_t^i}{2 + n_t} \\
\mathbf{P}([W_i = \text{true}] \mid [\mathit{Spam} = \text{true}]) = \dfrac{1 + n_t^i}{2 + n_t}
\end{cases}
\end{cases}
\end{cases} \\
\mathbf{Id} \rightarrow \text{Parameters learned from instances}
\end{cases} \\
\mathbf{Qu} \rightarrow \mathbf{P}(\mathit{Spam} \mid W_0 \wedge \ldots \wedge W_j \wedge \ldots \wedge W_{N-1})
\end{cases}
$$

Resultado



SpamSieve
http://c-command.com/spamsieve/

Objetos bayesianos

- Dominios:

$$[1.0, 270], \{1,2 \cdots 90\}, \{\text{Domingo, Lunes} \cdots \text{Sábado}\} \ldots$$

- Variables:

$$Distancia, Velocidad, Día, \ldots$$

- Valores de variables:

$$Distancia = 200, Velocidad = 120, Día = Martes$$

Objetos bayesianos

- Distribuciones condicionales

$$P(Distancia\ |Velocidad)$$

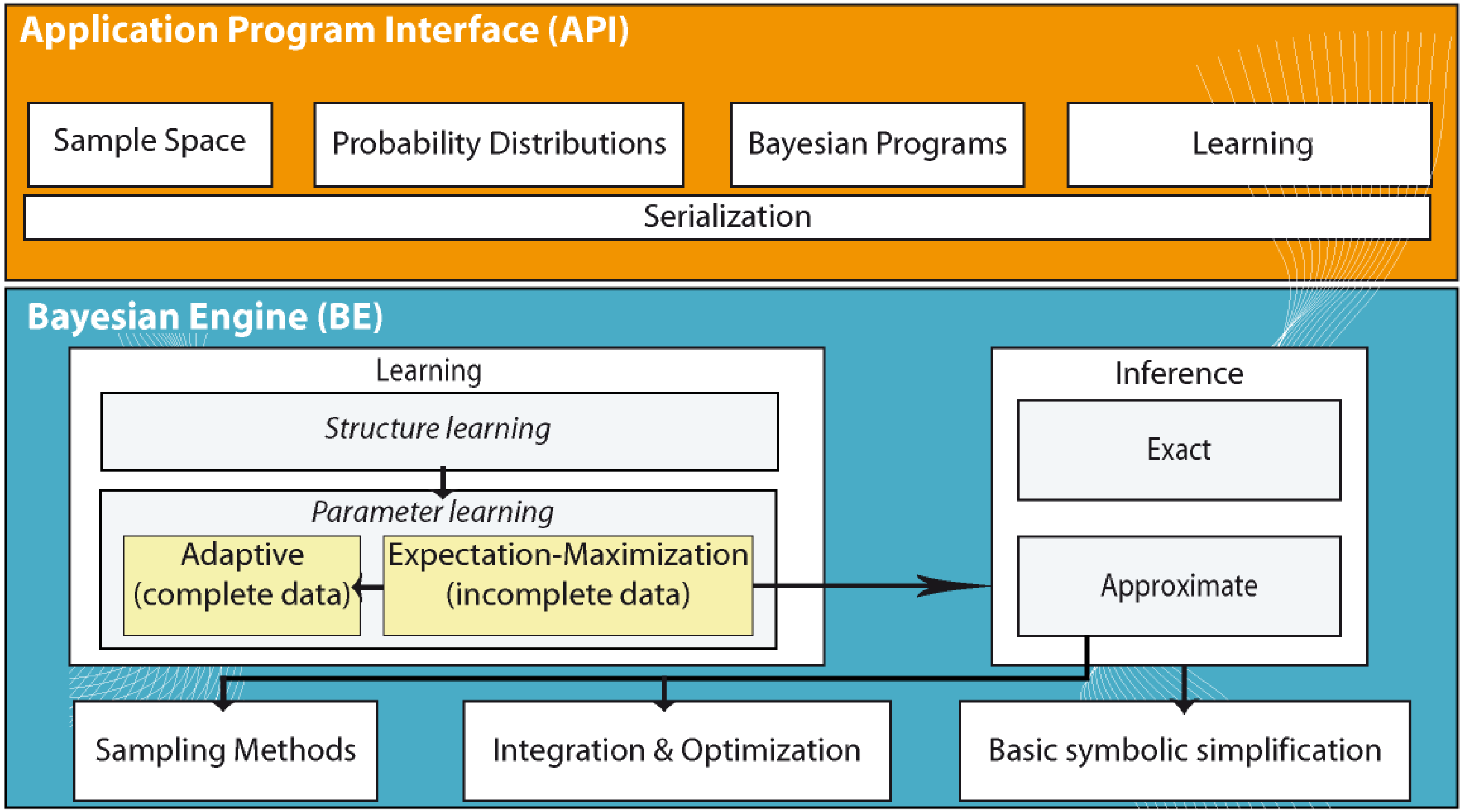
- Distribuciones no condicionales

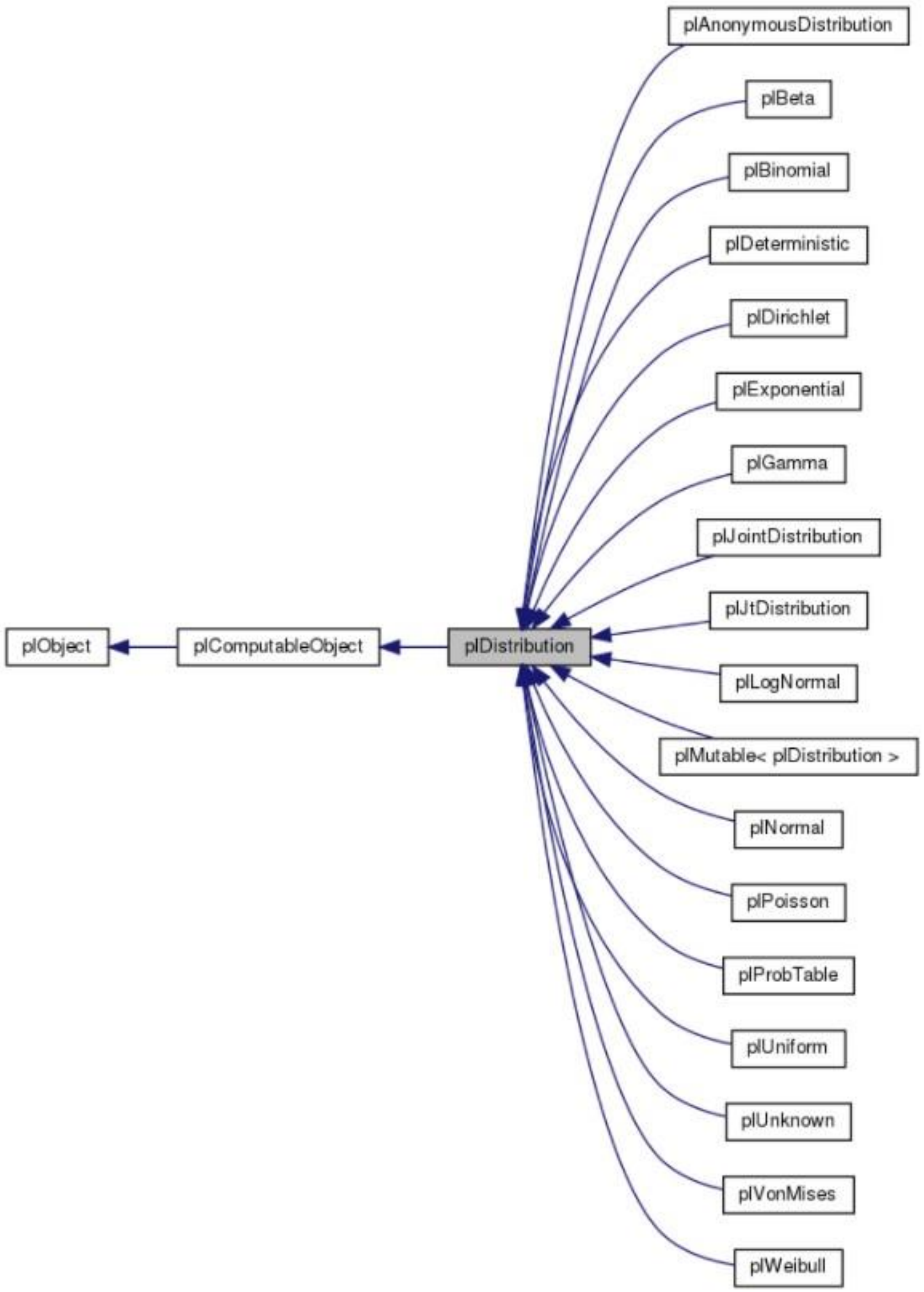
$$P(Distancia)$$
$$P(Distancia\ |Velocidad = 5)$$

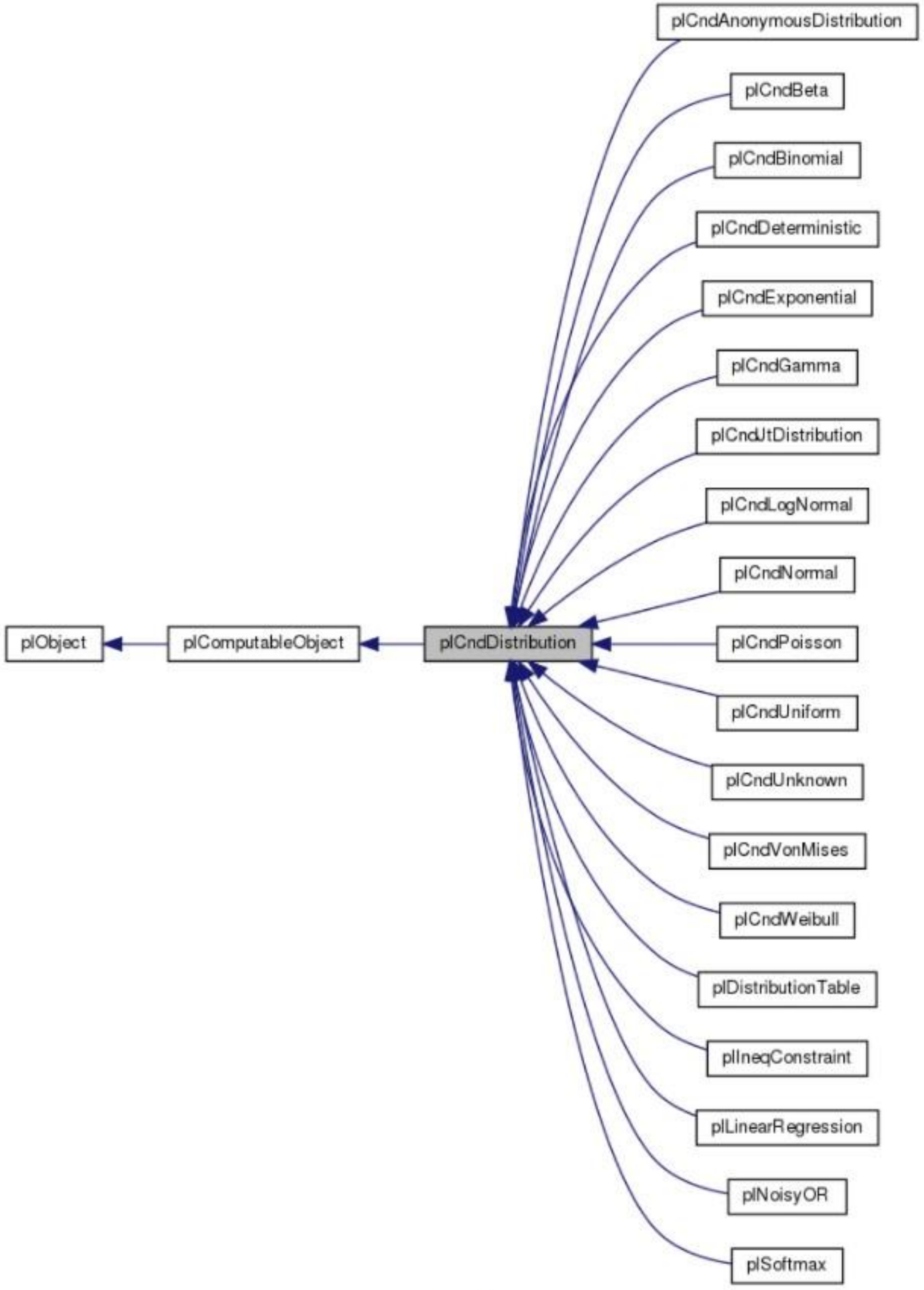
- Probabilidades

$$P(Distancia = 10)$$
$$P(Distancia = 30|Velocidad = 5)$$

Arquitectura de ProBT




plObject
plComputableObject
plDistribution
plAnonymousDistribution
plBeta
plBinomial
plDeterministic
plDirichlet
plExponential
plGamma
plJointDistribution
plJtDistribution
plLogNormal
plMutable< plDistribution >
plNormal
plPoisson
plProbTable
plUniform
plUnknown
plVonMises
plWeibull


plObject
plComputableObject
plCndDistribution
plCndAnonymousDistribution
plCndBeta
plCndBinomial
plCndDeterministic
plCndExponential
plCndGamma
plCndJtDistribution
plCndLogNormal
plCndNormal
plCndPoisson
plCndUniform
plCndUnknown
plCndVonMises
plCndWeibull
plDistributionTable
plIneqConstraint
plLinearRegression
plNoisyOR
plSoftmax

[legend]

Ejemplo 15: Tipos de variables

Ejemplo 16: Agrupación de variables

Operaciones con conjuntos

Producto cartesiano

El producto cartesiano de dos conjuntos A y B se define como el conjunto de todas las parejas ordenadas de (a,b) con a perteneciente a A y b perteneciente a B y se denota por " $\times$ "

$$A \times B = \{(a, b) : a \in A \text{ y } b \in B\}.$$

La cardinalidad de $\mathrm{A} \times B$ es $Card(a) * Card(b)$

Ejemplo 17: Iterar valores

$$P(Points) = Uniform$$

Ejemplo 18: Un dado

$$\mathrm{P(Points} \mid \mathrm{Die)} = \begin{cases} P_1(Die) & \text{si } Die = 1 \\ P_2(Die) & \text{si } Die = 2 \\ P_3(Die) & \text{si } Die = 3 \\ P_4(Die) & \text{si } Die = 4 \end{cases}$$

2.4 Modelos mixtos

Ejemplo 19: Dados diferentes

$BiasedDice() =$

- Description
  - Specification
    - Relevant Variables:
      $Die, Points$
    - Decomposition:
      $P(Die\ Points \mid \pi) =$
      $P(Die \mid \pi) P(Points \mid Die\ \pi)$
    - Parametric Forms:

$P(Die \mid \pi) =$

0	**1**
0.36	0.64

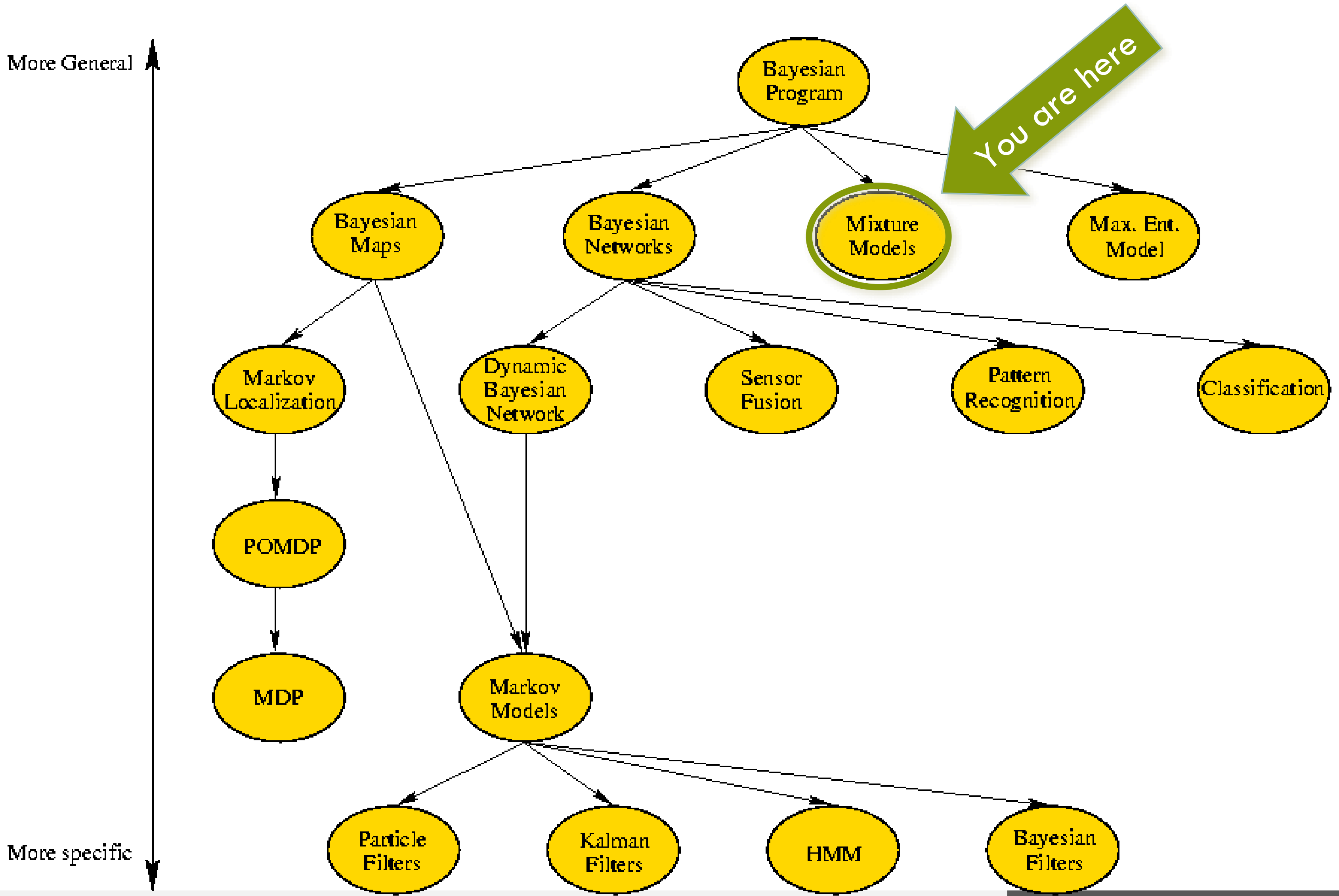
$P(Points \mid Die\ \pi) =$

Die	*Points*					
	1	**2**	**3**	**4**	**5**	**6**
0	0.3	0.2	0.1	0.1	0.2	0.1
1	0.16	0.16	0.16	0.16	0.16	0.16

  - Identification:
    All tables provided by the user
- Question:
  $P(Die|Points)$

Ejemplo 20: Dos dados

Jerarquía de los programas bayesianos



Actividad 05: El problema de los pigmeos

Proponer el programa bayesiano

https://docs.google.com/forms/d/e/1FAIpQLSciiGEJXKYRkee1YbZgf6Rs3oPOeA28jEegUbnygFT1H1nV6w/viewform?usp=sf_link

El problema de los pigmeos

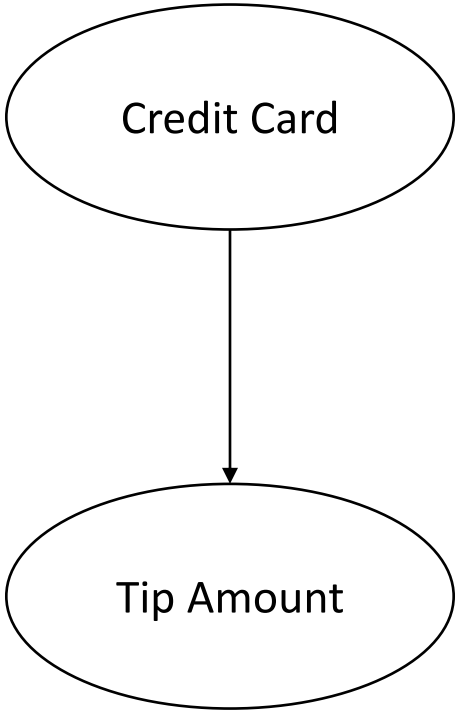
La población de una región del África ecuatorial está compuesta por una mayoría no pigmea (84%) y una minoría pigmea (16%). Un adulto no pigmeo tiene una altura media de 1.55 metros con una desviación estándar de 7.5 cm. En cambio, un adulto pigmeo tiene una altura media de 1.30 metros con una desviación estándar de 9.5 cm. Se supone que la distribución de la altura es gaussiana.

Propón el programa bayesiano para calcular P(Altura) para esta región de África Ecuatorial.

Un modelo para los datos de ventas

- Program
  - Description
    - *Specification*
      - Variables
      - Decomposition
      - Parametrical Forms or Recursive Question
    - *Identification*
  - Question
    - Use

Modelo de tarjeta de crédito y monto de propina



Modelo de tarjeta de crédito y monto de propina

Program

Description

Specification

- Variables

$CreditCard \in \{False, True\}$

$TipAmount \in [0, 100]$

- Decomposition

$P(CreditCard\ TipAmount) = P(CreditCard)P(TipAmount\ |CreditCard)$

- Parametrical Forms or Recursive Question

$P(CreditCard)$ = ProbabilityTable

$P(TipAmount\ |CreditCard) = \text{collection of Normal distributions}$

Identification

Data from file `TaxiTripSmall.xlsx`

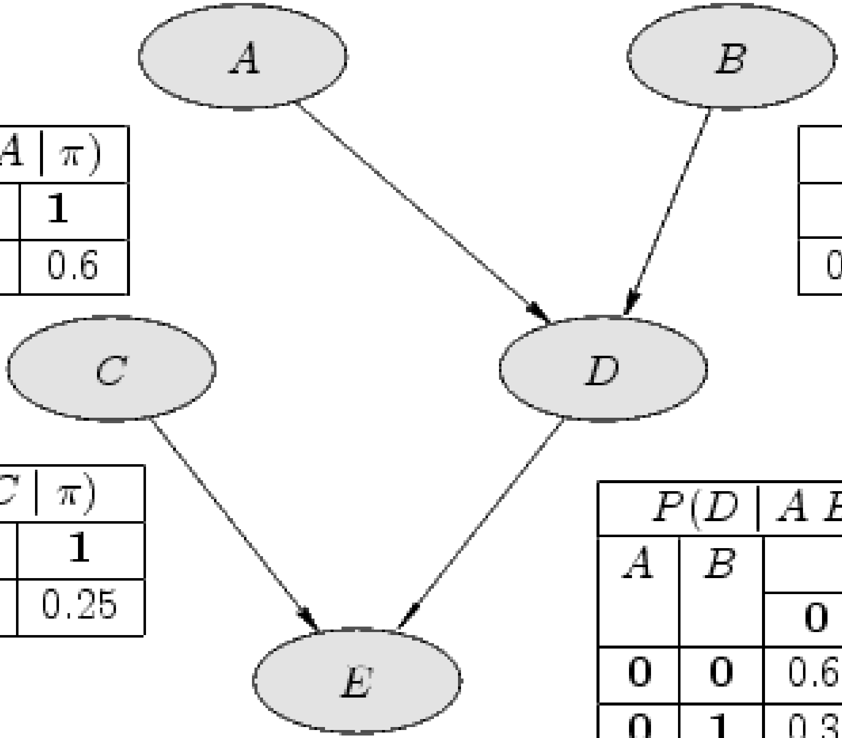
Question

$P(CreditCard\ |TipAmmount)$

Ejemplo 21: Resolviendo el ejemplo 14 con ProBT

Creación de equipos

Ejemplo 22: Red bayesiana



$P(A \mid \pi)$	
0	**1**
0.4	0.6

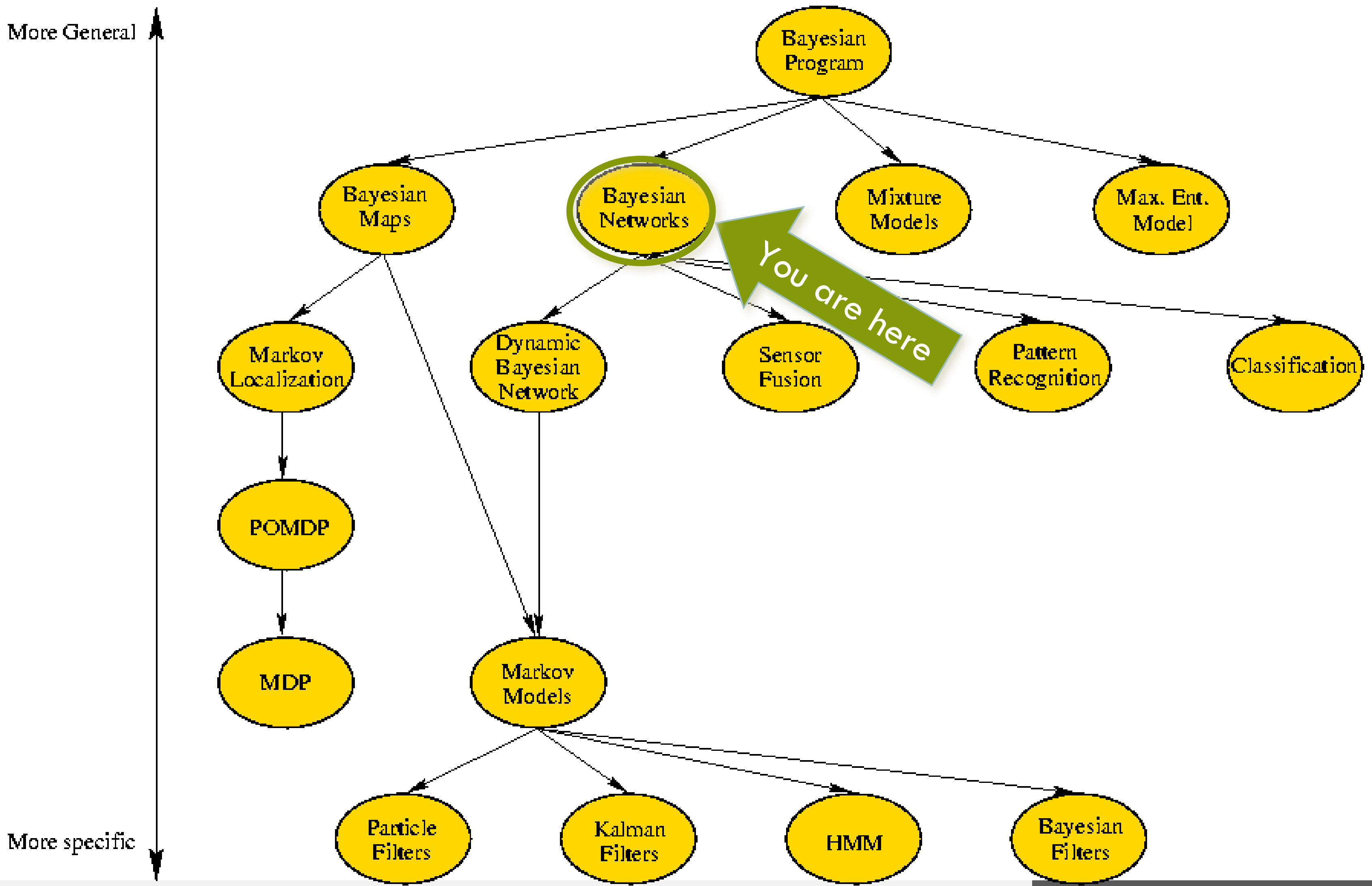
$P(B \mid \pi)$	
0	**1**
0.18	0.82

$P(C \mid \pi)$	
0	**1**
0.75	0.25

$P(D \mid A\ B\ \pi)$			
A	B	D	
		0	**1**
0	**0**	0.6	0.4
0	**1**	0.3	0.7
1	**0**	0.1	0.9
1	**1**	0.5	0.5

$P(E \mid C\ D\ \pi)$			
C	D	WG	
		0	**1**
0	**0**	0.59	0.41
0	**1**	0.25	0.75
1	**0**	0.8	0.2
1	**1**	0.35	0.65

Jerarquía de los programas bayesianos



Ejercicio 07: Descomposición

Obtener la descomposición a partir de la red bayesiana

https://docs.google.com/forms/d/e/1FAIpQLSdWMvgXS-6qEs2mWFor-CqdZf6xbyMBZ3jJAc_i8BpXcXYedw/viewform?usp=sf_link

Actividad 06:
Una situación complicada para Peter

Identificar las variables

1.	RI	Rain on invitation's day	Ales
2.	RP	Rain on Party's day	
3.	We	Invitation for aWeekend	
4.	LI	Late invitation	
5.	E	Emergency	
6.	JA	John accepted	Acceptation
7.	MA	Mary accepted	
8.	AA	Alice accepted	
9.	BA	Bill accepted	
10.	JP	John is present	Presence
11.	MP	Mary is present	
12.	AP	Alice is present	
13.	VP	Victor is present	
14.	BP	Bill is present	

Escribir la distribución conjunta

La especificación del programa bayesiano

$$
PartyBN() = \begin{cases}
\text{Description} \begin{cases}
\text{Specification} \begin{cases}
\text{Relevant Variables:} \\
\quad RI, RP, We, LI, E, JA, MA, AA, BA, JP, MP, AP, VP, BP \\
\text{Decomposition:} \\
P(RI\ RP\ We\ LI\ E\ JA\ MA\ AA\ BA\ JP\ MP\ AP\ VP\ BP \mid \pi) = \\
\qquad P(RI \mid \pi) P(RP \mid \pi) P(We \mid \pi) P(LI \mid \pi) P(E \mid \pi) \\
\qquad P(JA \mid RI\ \pi) P(BA \mid LI\ \pi) P(AA \mid We\ \pi) P(AP \mid AA\ \pi) \\
\qquad P(MA \mid JA\ AA\ \pi) P(MP \mid MA\ \pi) P(VP \mid AA\ BA\ \pi) \\
\qquad P(JP \mid JA\ RP\ \pi) P(BP \mid BA\ E\ \pi) \\
\text{Parametric Forms:} \\
\text{See Figure3}
\end{cases} \\
\text{Identification:} \\
\quad \text{All tables provided by the user}
\end{cases} \\
\text{Question:} \\
\quad P(RI \mid MP\ We\ \pi)
\end{cases}
$$

Escribir las formas paramétricas

$P(RI \mid \pi)$	
0	**1**

$P(RP \mid \pi)$	
0	**1**

$P(We \mid \pi)$	
0	**1**

$P(LI \mid \pi)$	
0	**1**

$P(E \mid \pi)$	
0	**1**

$P(JA \mid RI\ \pi)$		
RI	JA	
	0	**1**
0		
1		

$P(BA \mid LI\ \pi)$		
LI	BA	
	0	**1**
0		
1		

$P(AA \mid We\ \pi)$		
We	AA	
	0	**1**
0		
1		

$P(AP \mid AA\ \pi)$		
AA	AP	
	0	**1**
0		
1		

$P(MA \mid JA\ AA\ \pi)$			
JA	AA	MA	
		0	**1**
0	**0**		
0	**1**		
1	**0**		
1	**1**		

$P(MP \mid MA\ \pi)$		
MA	MP	
	0	**1**
0		
1		

$P(VP \mid AA\ BA\ \pi)$			
AA	BA	VP	
		0	**1**
0	**0**		
0	**1**		
1	**0**		
1	**1**		

$P(JP \mid JA\ RP\ \pi)$			
JA	RP	JP	
		0	**1**
0	**0**		
0	**1**		
1	**0**		
1	**1**		

$P(BP \mid BA\ E\ \pi)$			
BA	E	BP	
		0	**1**
0	**0**		
0	**1**		
1	**0**		
1	**1**		

Escribir las formas paramétricas

$P(RI \mid \pi)$	
0	**1**
0.78	0.22

$P(RP \mid \pi)$	
0	**1**
0.78	0.22

$P(We \mid \pi)$	
0	**1**
0.72	0.28

$P(LI \mid \pi)$	
0	**1**
0.7	0.3

$P(E \mid \pi)$	
0	**1**
0.35	0.65

$P(JA \mid RI\ \pi)$		
RI	JA	
	0	**1**
0	0.1	0.9
1	0.6	0.4

$P(BA \mid LI\ \pi)$		
LI	BA	
	0	**1**
0	0.05	0.95
1	0.7	0.3

$P(AA \mid We\ \pi)$		
We	AA	
	0	**1**
0	0.6	0.4
1	0.2	0.8

$P(AP \mid AA\ \pi)$		
AA	AP	
	0	**1**
0	0.5	0.5
1	0.2	0.8

$P(MA \mid JA\ AA\ \pi)$			
JA	AA	MA	
		0	**1**
0	**0**	0.5	0.5
0	**1**	0.15	0.85
1	**0**	0.05	0.95
1	**1**	0.95	0.05

$P(MP \mid MA\ \pi)$		
MA	MP	
	0	**1**
0	0.97	0.03
1	0.03	0.97

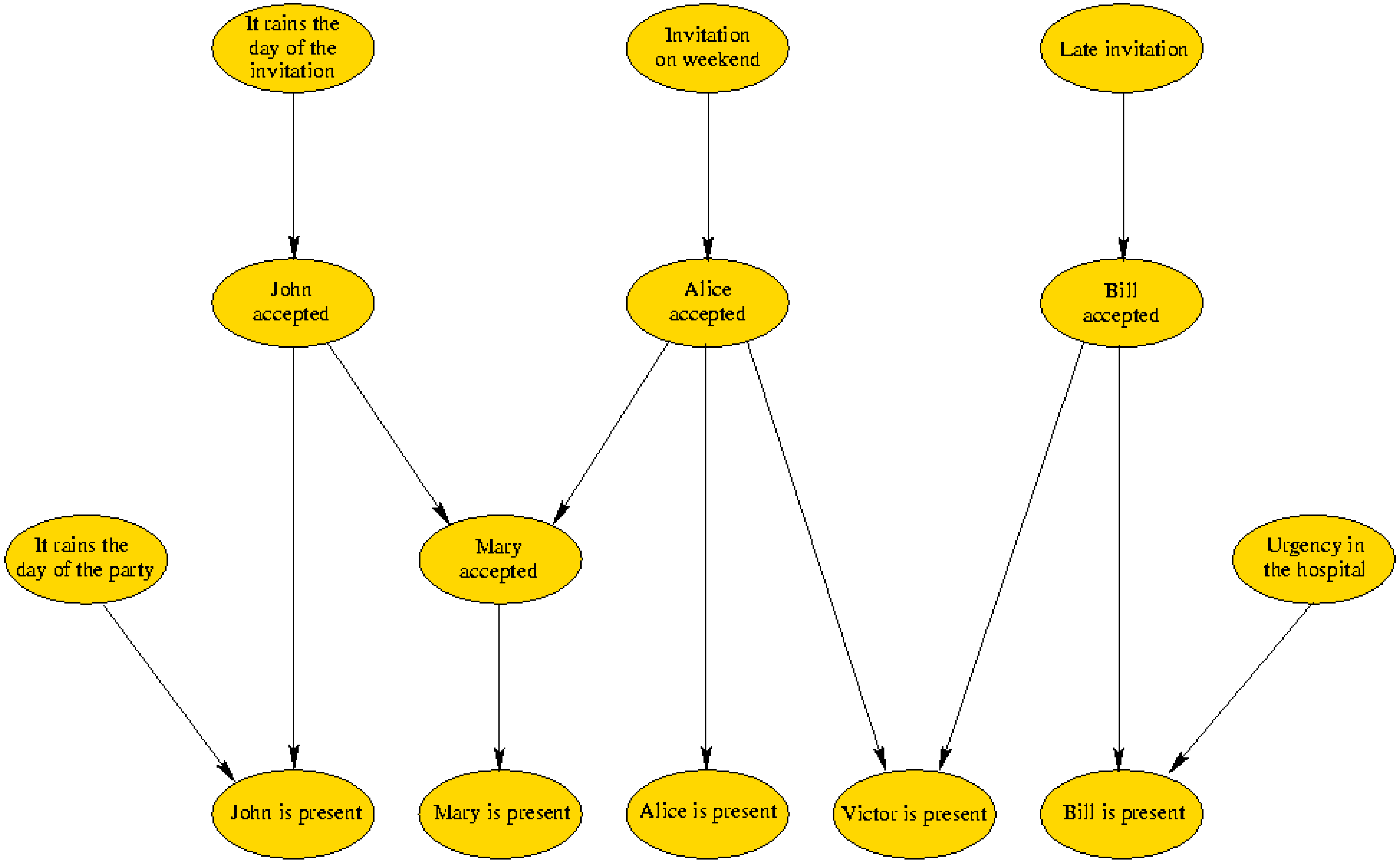
$P(VP \mid AA\ BA\ \pi)$			
AA	BA	VP	
		0	**1**
0	**0**	0.3	0.7
0	**1**	0.6	0.4
1	**0**	0.1	0.9
1	**1**	0.0	1.0

$P(JP \mid JA\ RP\ \pi)$			
JA	RP	JP	
		0	**1**
0	**0**	0.6	0.4
0	**1**	1.0	0.0
1	**0**	0.1	0.9
1	**1**	0.3	0.7

$P(BP \mid BA\ E\ \pi)$			
BA	E	BP	
		0	**1**
0	**0**	0.8	0.2
0	**1**	1.0	0.0
1	**0**	0.05	0.95
1	**1**	1.0	0.0

Dibujar la red bayesiana

La red bayesiana



Escribir la expression para las siguientes preguntas

1.- What is the probability that it was raining when John received the invitation knowing that Mary didn't attend the party and that the party is on Monday?

2.- What is the probability that all Peter's friends attend the party knowing that it will take place on a sunny Saturday, knowing that Bill received the invitation in time and no emergency was present at the hospital?

3.- What is the probability that Alice, Victor and Bill attend the party?

4.- Alice answer the phone at Peter's house, "Hi John". What is the probability that Mary and Victor attended the party?

$$P(Points_1 Points_2\ Sum) = P(Points_1)P(Points_2)P(Sum\ |Points_1 Points_2)$$

Funciones externas

$$\mathrm{P}(Points_1 Points_2\ Sum) = P(Points_1)P(Points_2)P(Sum\ |Points_1 Points_2)$$

$$P(Points_1) = \mathrm{Uniform}(Points_1)$$

$$P(Points_2) = \mathrm{Uniform}(Points_2)$$

$$P(Sum\ |Points_1 Points_2) = \begin{cases} 1 & \text{si } Sum = Points_1 + Points_2 \\ 0 & \text{de otra forma} \end{cases}$$

Funciones externas

$$P(S\,|K) = \begin{cases} 1 & \text{si } S = f(K) \\ 0 & \text{de otra form}a \end{cases}$$

$f(K)$ se dice que es la **función externa.**
$P(S\,|K)$ se dice que es una **distribución condicional determinística.**

Ejemplo 23: Suma de puntos

Datos de ventas

	Date	Day_name	Day	Pay_day	Holiday	Rain	Promotion	Visits	Sales	ASPV
0	2017-01-01	Saturday	6	1	1	0	0	1403	127879.1	91.1469
1	2017-01-02	Sunday	0	1	0	0	0	1809	152746	84.43669
2	2017-01-03	Monday	1	0	0	0	0	788	43802.62	55.58709
3	2017-01-04	Tusday	2	0	0	1	0	410	48217.33	117.6032
4	2017-01-05	Wednesday	3	0	0	0	0	766	59778.45	78.03975
5	2017-01-06	Thursday	4	0	0	0	0	717	74412.95	103.7837
6	2017-01-07	Friday	5	0	0	0	0	911	127786	140.27
7	2017-01-08	Saturday	6	0	0	0	0	1093	125859.7	115.1507
8	2017-01-09	Sunday	0	0	0	1	0	865	119650.2	138.3239
9	2017-01-10	Monday	1	0	0	1	0	313	34795.89	111.169
10	2017-01-11	Tusday	2	0	0	1	0	300	39124.33	130.4144
11	2017-01-12	Wednesday	3	0	0	0	0	808	70751.65	87.56392
12	2017-01-13	Thursday	4	0	0	1	0	472	58457.66	123.851
13	2017-01-14	Friday	5	1	0	0	0	1164	77431.53	66.52194
14	2017-01-15	Saturday	6	1	0	1	0	720	126103.4	175.1436
15	2017-01-16	Sunday	0	1	0	0	0	1644	184111.3	111.9898
16	2017-01-17	Monday	1	0	0	0	0	703	39910.77	56.77208

Un modelo para los datos de ventas

- ¿Qué variables son independientes?
- ¿Qué variables son dependientes?
- ¿Cuáles son los tipos de las variables?
- ¿Cuál es la pregunta que nos importa?

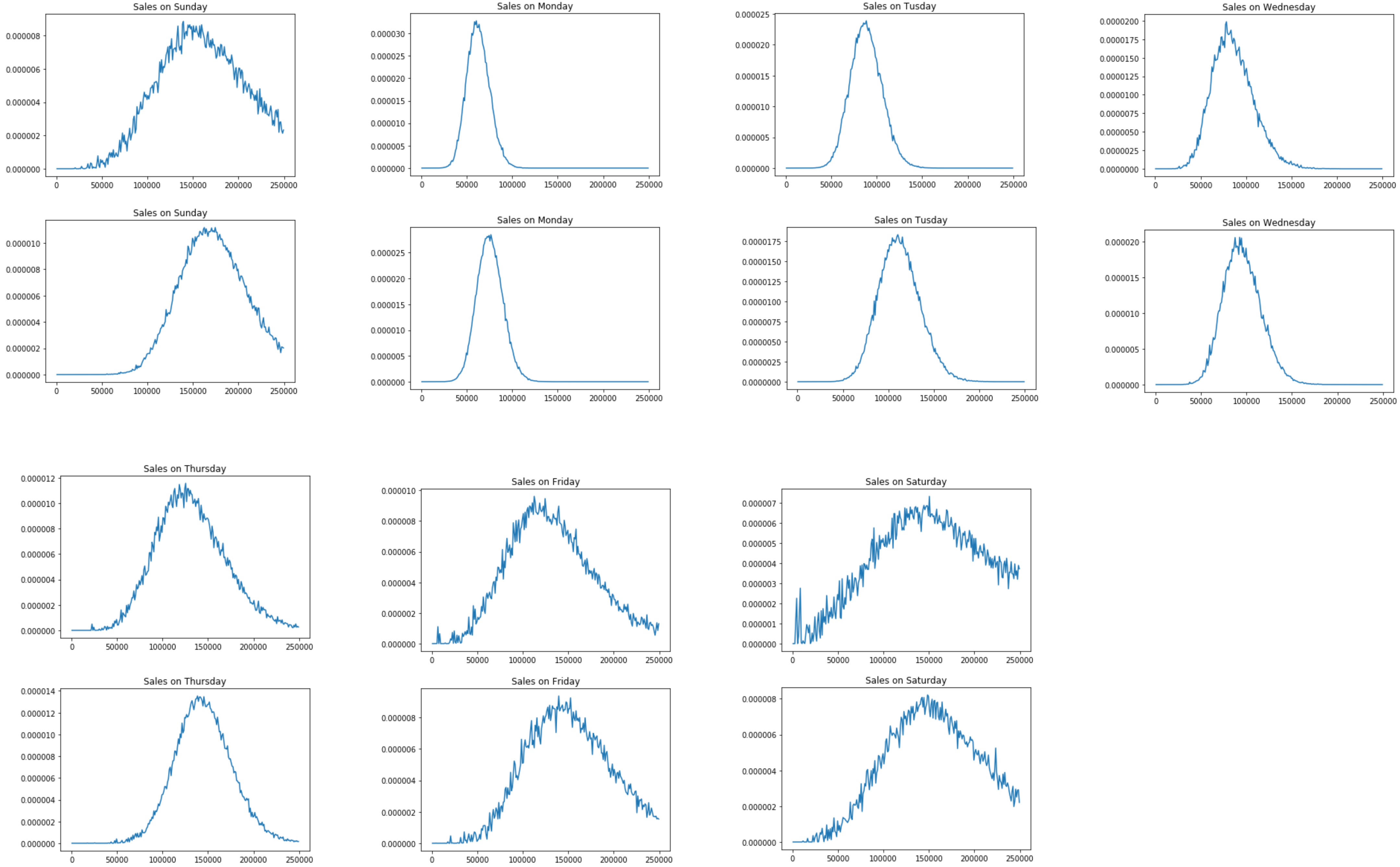
Un modelo para los datos de ventas

- Program
  - Description
    - *Specification*
      - Variables
      - Decomposition
      - Parametrical Forms or Recursive Question
    - *Identification*
  - Question
    - Use

Ejercicio 08: Definir el Programa bayesiano para los datos de ventas

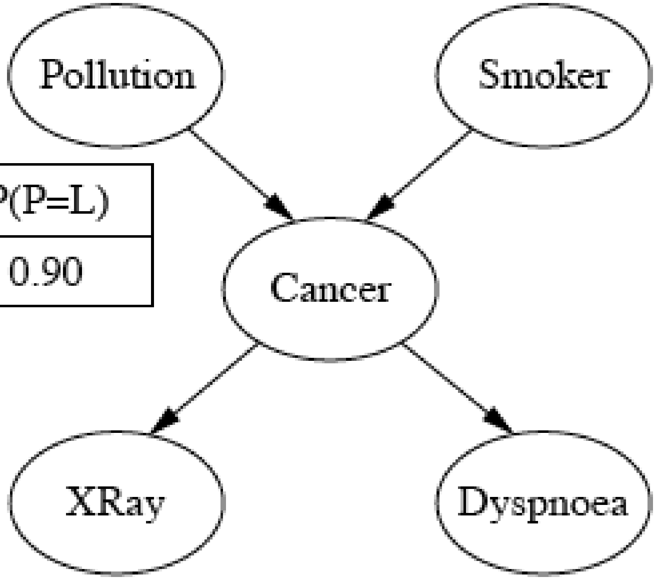
https://docs.google.com/forms/d/e/1FAIpQLSdqLo0D-foy7Alif8XNZH3EKN4GfrckRdSV9RuZvoxMJKJ6EA/viewform?usp=sf_link

Ejemplo 24: Modelo de ventas


Sales on Sunday
Sales on Monday
Sales on Tusday
Sales on Wednesday
Sales on Sunday
Sales on Monday
Sales on Tusday
Sales on Wednesday
Sales on Thursday
Sales on Friday
Sales on Saturday
Sales on Thursday
Sales on Friday
Sales on Saturday

Otros ejemplos

Node name	Type	Values
Pollution	Binary	{*low*, *high*}
Smoker	Boolean	{*T*, *F*}
Cancer	Boolean	{*T*, *F*}
Dyspnoea	Boolean	{*T*, *F*}
X-ray	Binary	{*pos*, *neg*}



P(S=T)
0.30

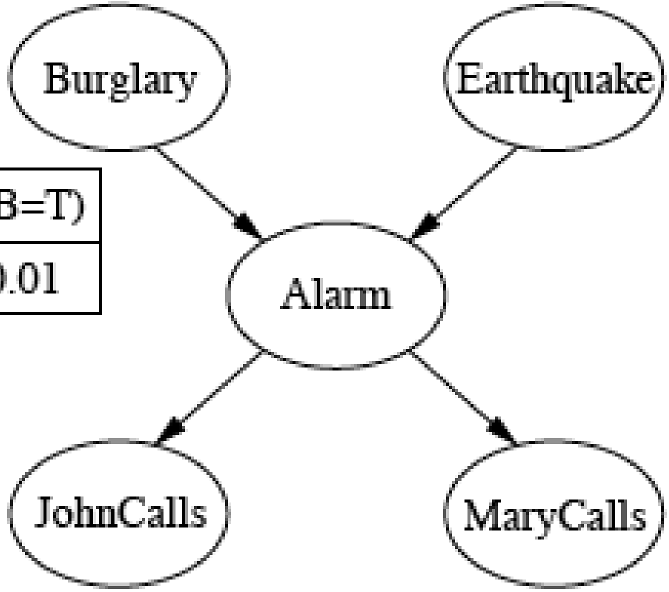
P(P=L)
0.90

P	S	P(C=T|P,S)
H	T	0.05
H	F	0.02
L	T	0.03
L	F	0.001

C	P(X=pos|C)
T	0.90
F	0.20

C	P(D=T|C)
T	0.65
F	0.30

Otros ejemplos



P(B=T)
0.01

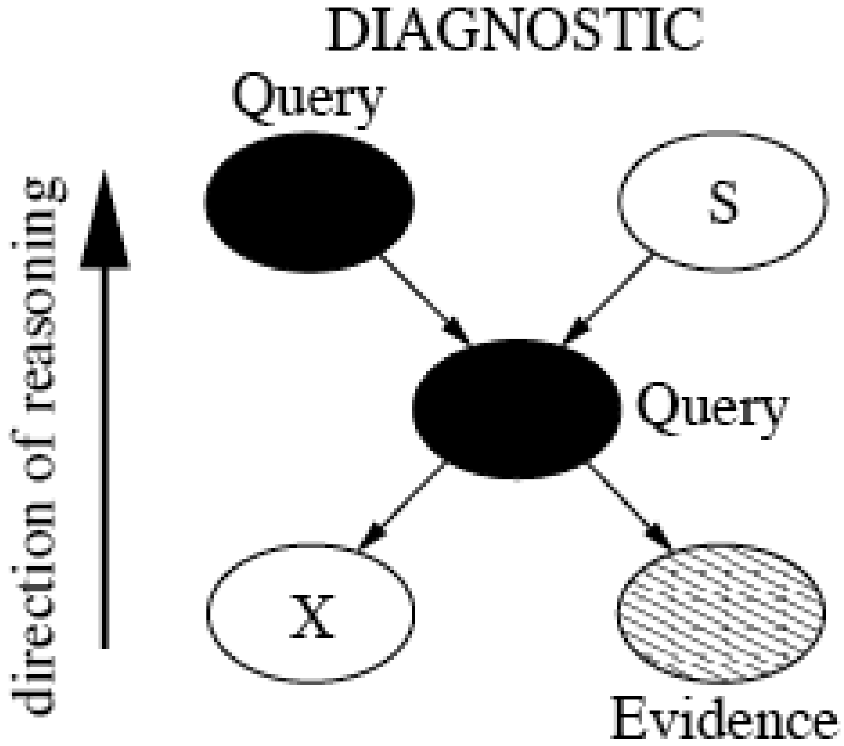
P(E=T)
0.02

B	E	P(A=T|B,E)
T	T	0.95
T	F	0.94
F	T	0.29
F	F	0.001

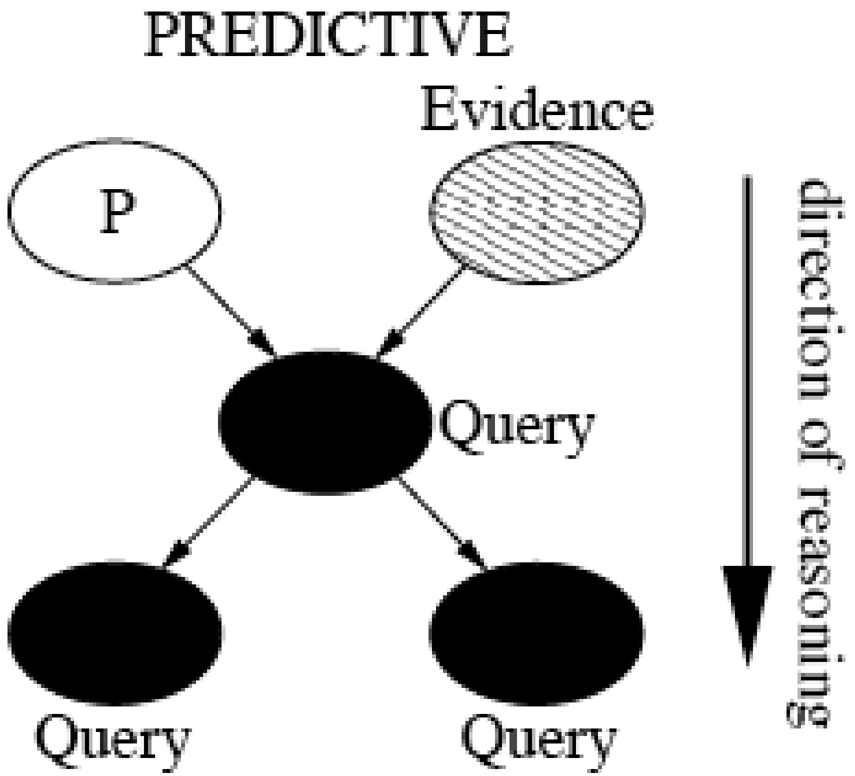
A	P(J=T|A)
T	0.90
F	0.05

A	P(M=T|A)
T	0.70
F	0.01

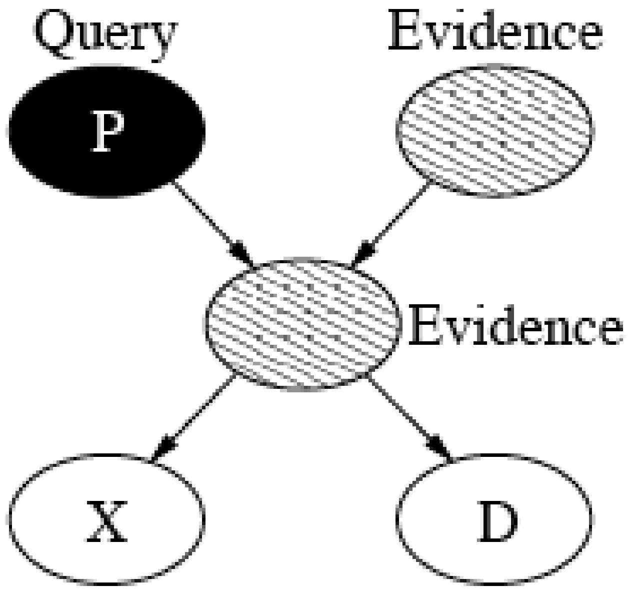
Diagnóstico



Predicción

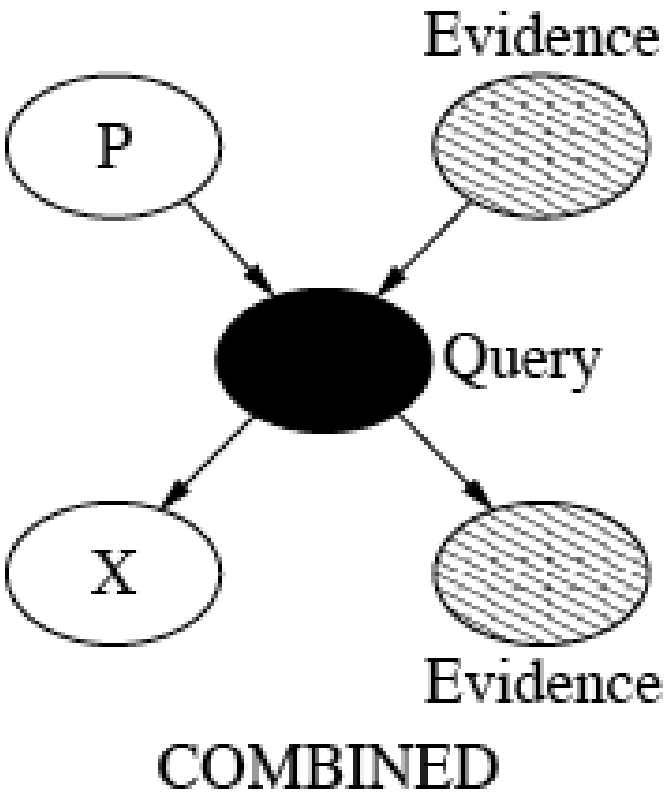

PREDICTIVE
Evidence
P
Query
Query
Query
direction of reasoning

Intercasual

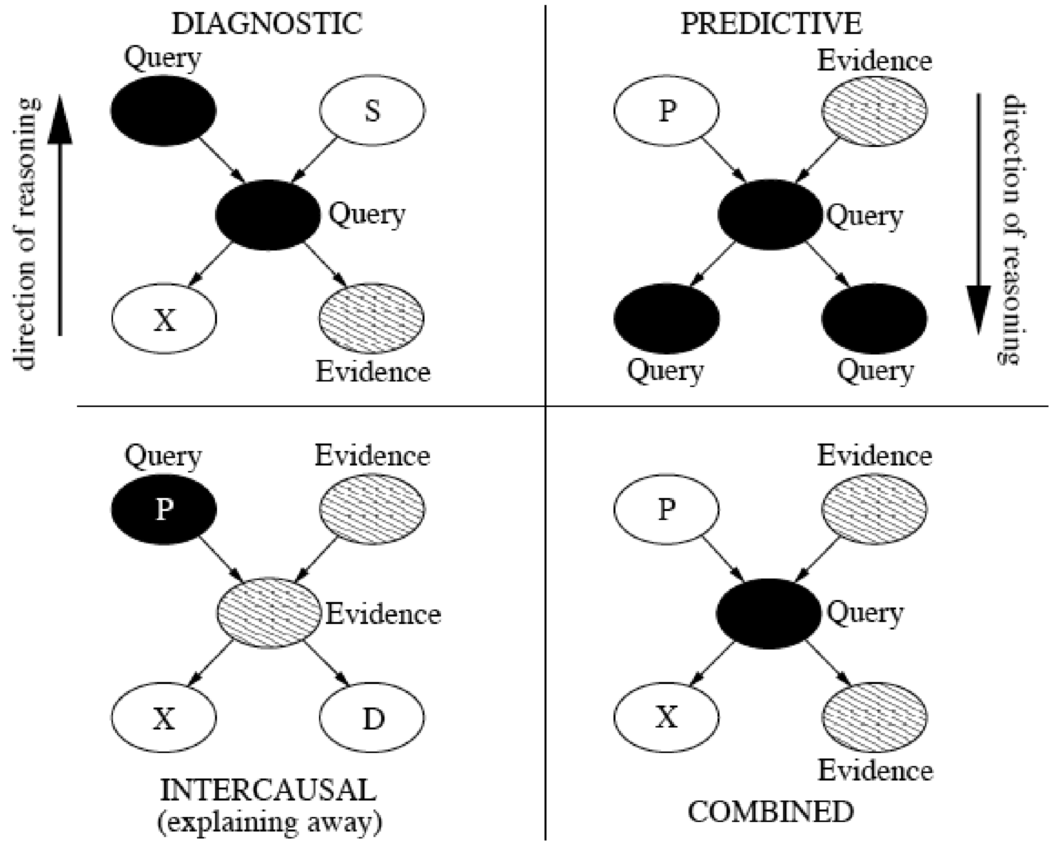


INTERCAUSAL
(explaining away)

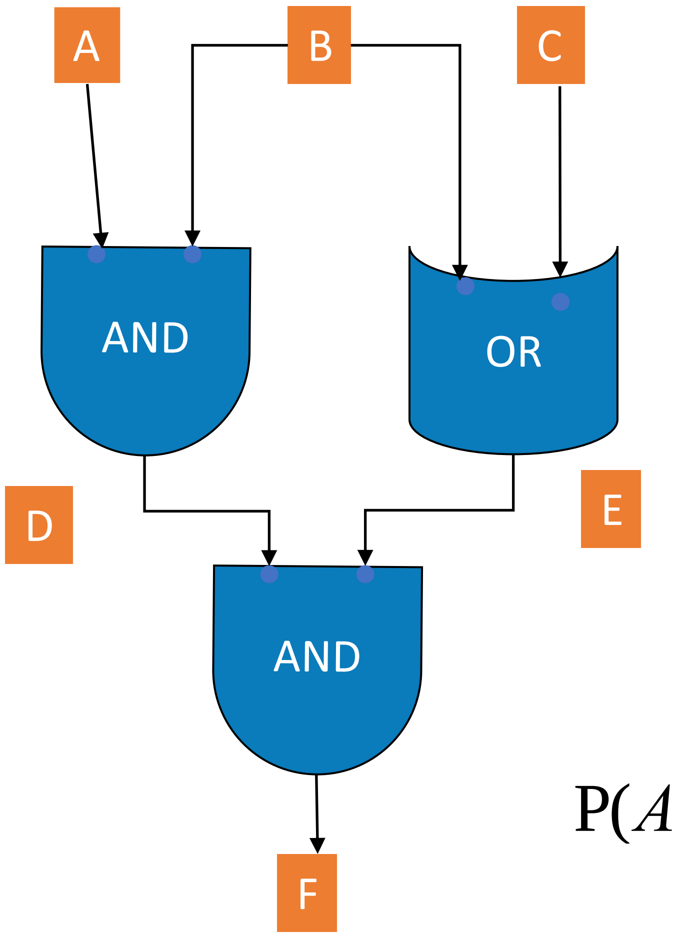
Combinada

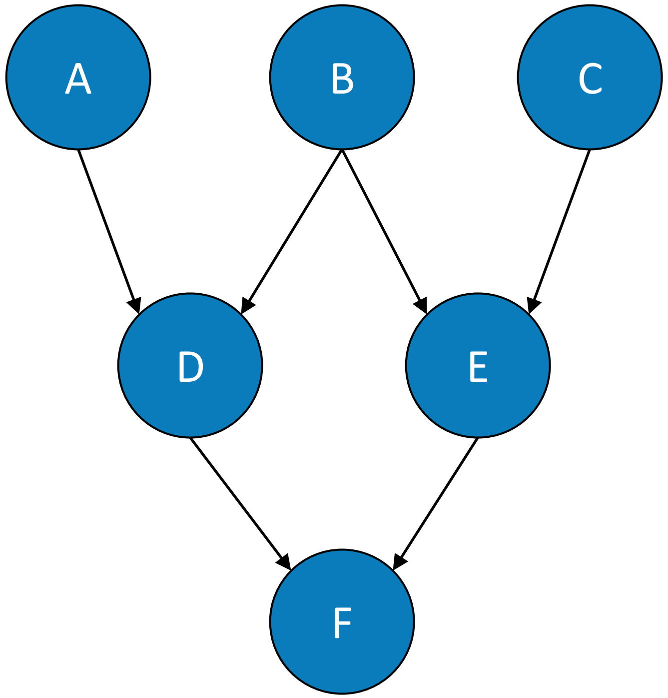


Tipos de razonamiento o inferencia



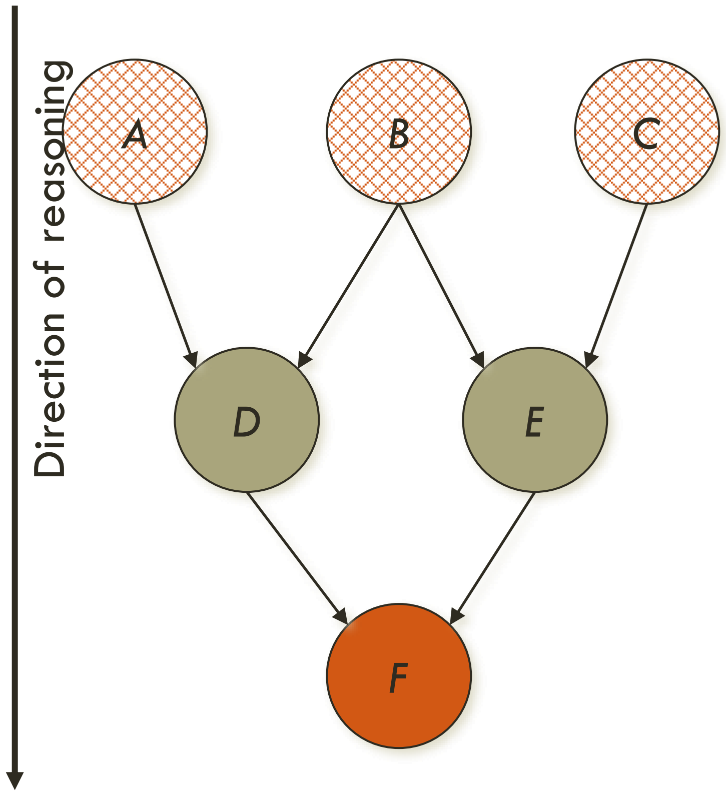
Tipos de razonamiento o inferencia



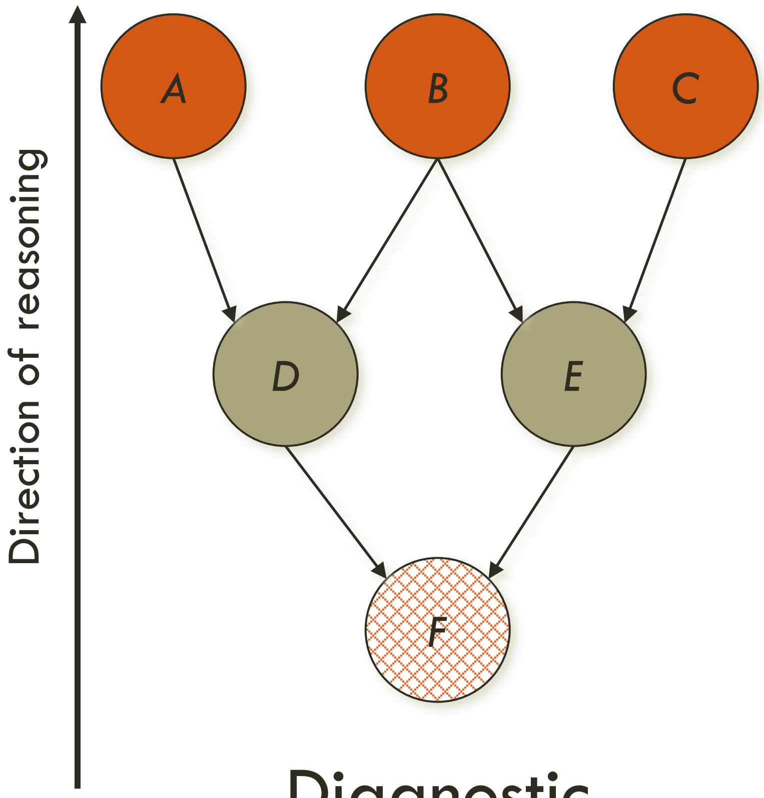


$$P(A)P(B)P(C)P(D \mid A\ B)P(E \mid B\ C)P(F \mid D\ E)$$

Tipos de razonamiento o inferencia



Predictive



Diagnostic

Tipos de razonamiento o inferencia

Direction of reasoning

Predictive

$$P(a)P(b)P(c)P(D \mid A\,B)P(E \mid B\,C)P(F \mid D\,E)$$

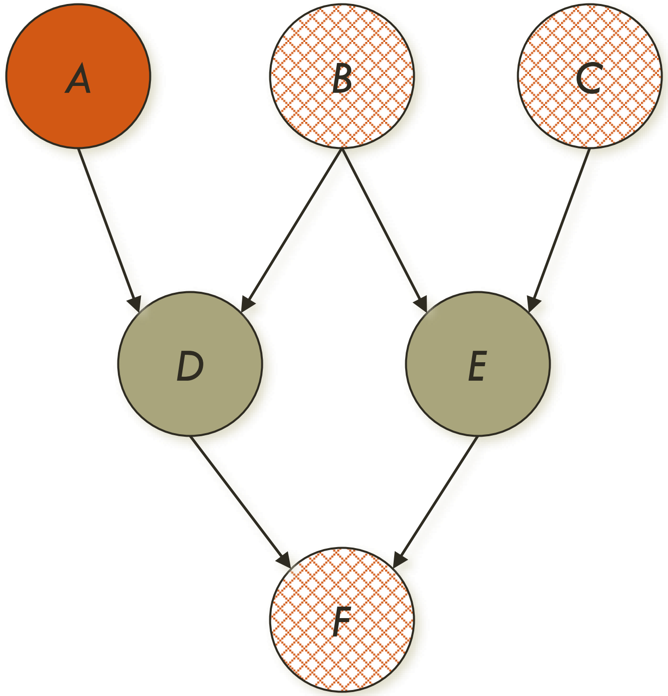
Direction of reasoning

Diagnostic

$$P(A)P(B)P(C)P(D \mid A\,B)P(E \mid B\,C)P(f \mid D\,E)$$

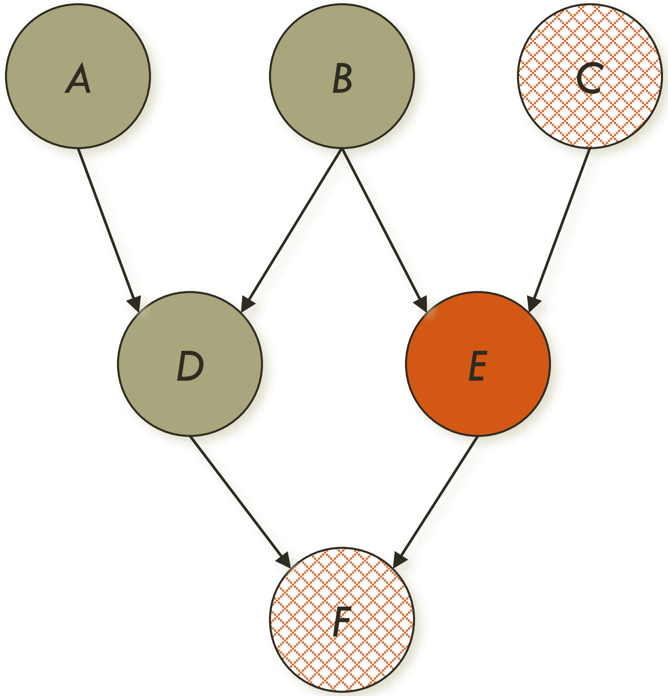
Tipos de razonamiento o inferencia

$$P(A \mid B\, C\, F)$$



Intercausal

$$P(E \mid C\, F)$$



Mixed

Tipos de razonamiento o inferencia

Intercausal

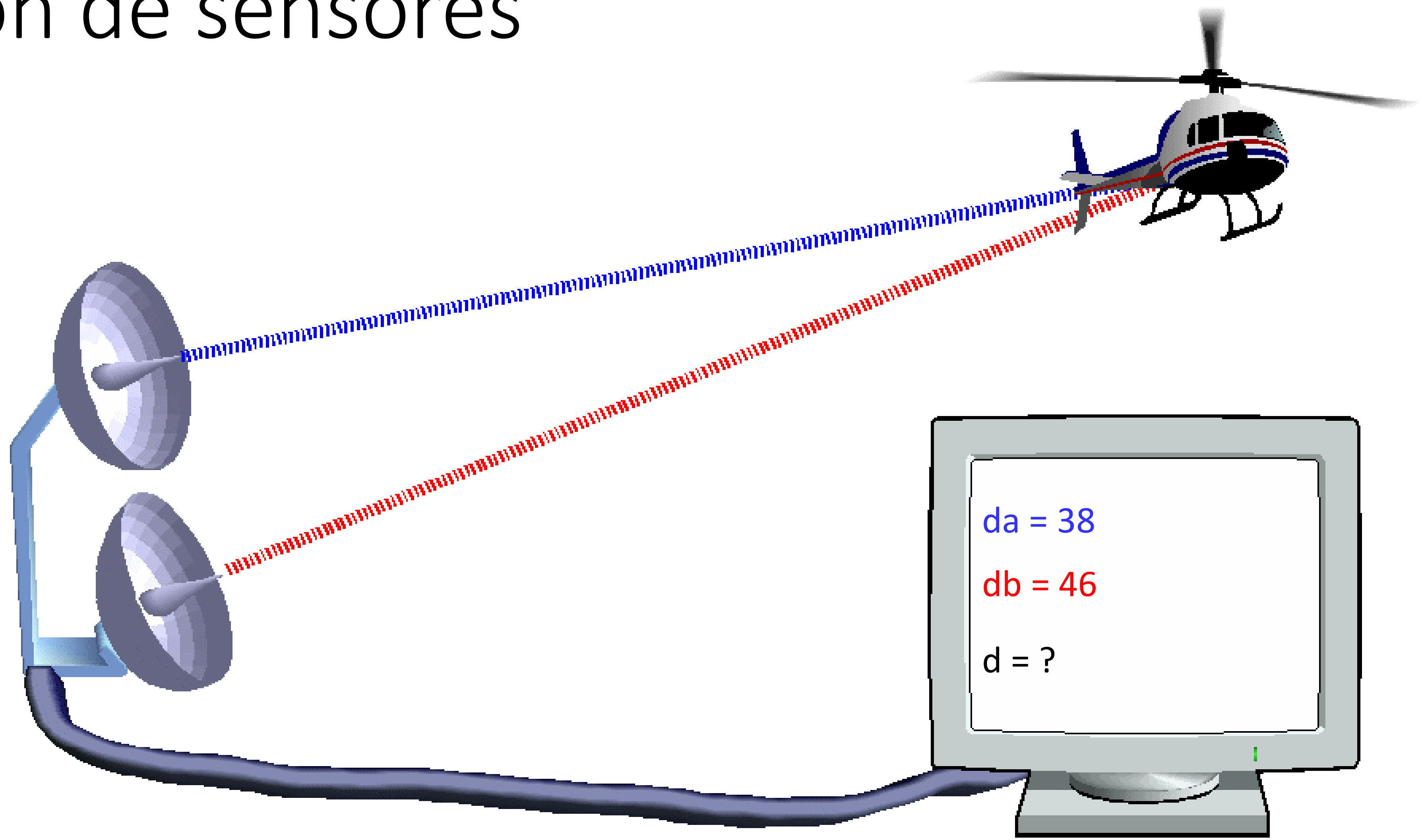
$$\mathrm{P}(A)\mathrm{P}(b)\mathrm{P}(c)\mathrm{P}(D \mid A\, b)\mathrm{P}(E \mid b\, c)\mathrm{P}(f \mid D\, E)$$

Mixed

$$\mathrm{P}(A)\mathrm{P}(B)\mathrm{P}(c)\mathrm{P}(d \mid A\, B)\mathrm{P}(E \mid B\, c)\mathrm{P}(f \mid d\, E)$$

2.5 Fusión de sensores

Fusión de sensores



Fusión de sensores

Identificación de variables

$$d \in [1, 100]$$
$$da \in [1, Max_a]$$
$$db \in [1, Max_b]$$

Variables:

d,

da,

db.

Types:

$t1 = [1,100]$

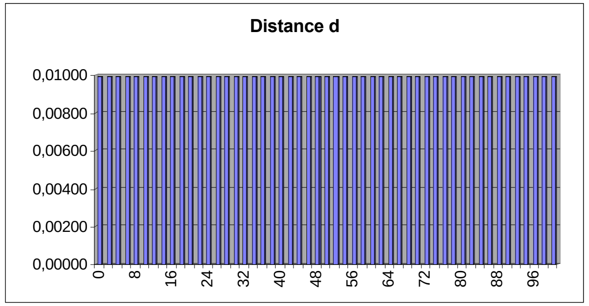
$t2 = [1, Max_a]$

$t3 = [1, Max_b]$

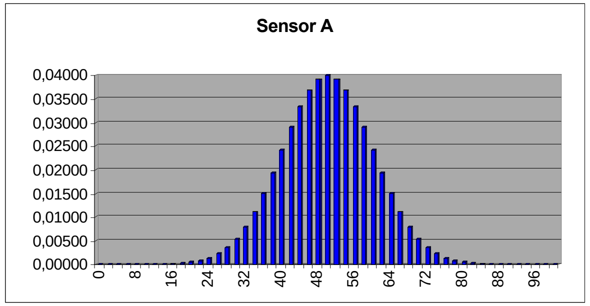
Fusión de sensores

Formas paramétricas

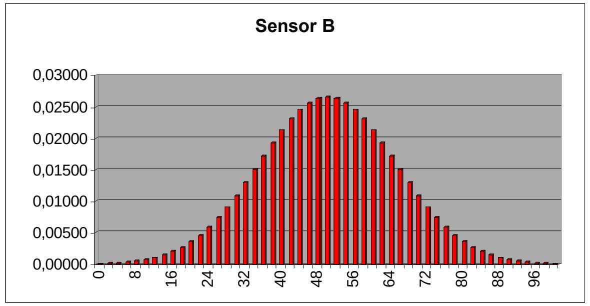
P(d) = uniform(d)



P(da | d) = CndBellShape(da, d,10)



P(db | d)= CndBellShape(da, d, 15)



Fusión de sensores

Distribución conjunta y pregunta

$$P(d\ da\ db) = P(d)\ P(da \mid d)\ P(db \mid d)$$

Pregunta

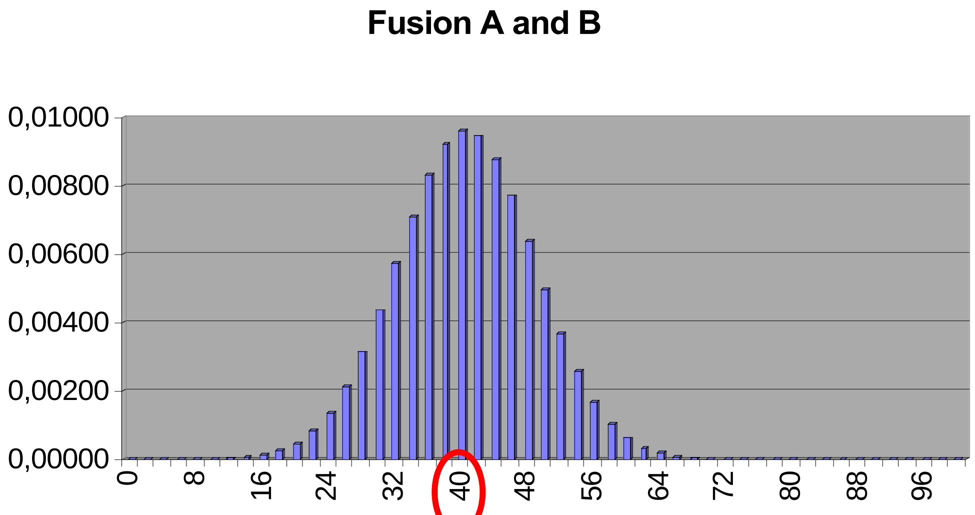
$$Best(P(d \mid da=va\ db=vb))$$

Fusión de sensores

Observaciones

1.- P(d | da db) es una función de tres variables.

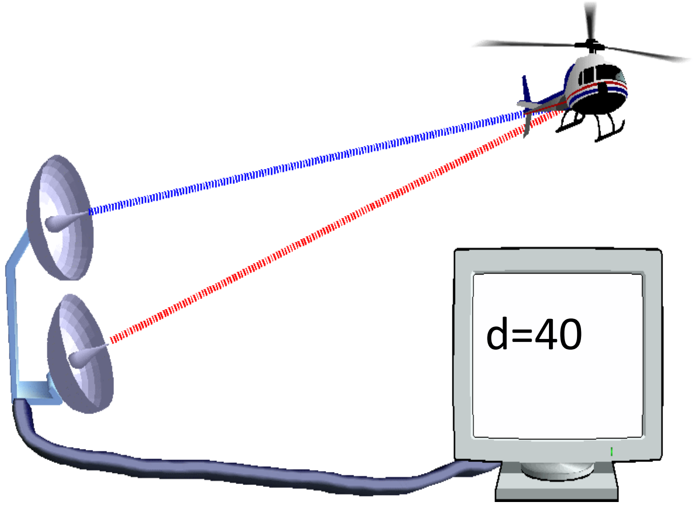
2.- P(d | da=va db=vb) es una función de una variable.



3.- Best(P(d | da=va db=vb)) es un escalar.

Ejemplo 25: Fusión de 2 sensores

Fusión de sensores



Distribución conjunta

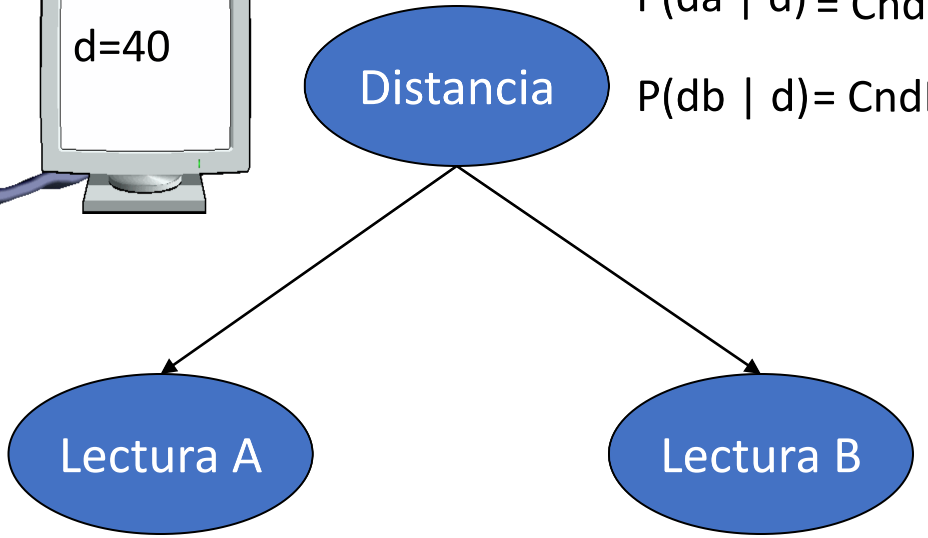
$$P(d\ da\ db) = P(d)\ P(da \mid d)\ P(db \mid d)$$

Formas paramétricas:

$P(d) = \text{uniform}(d)$

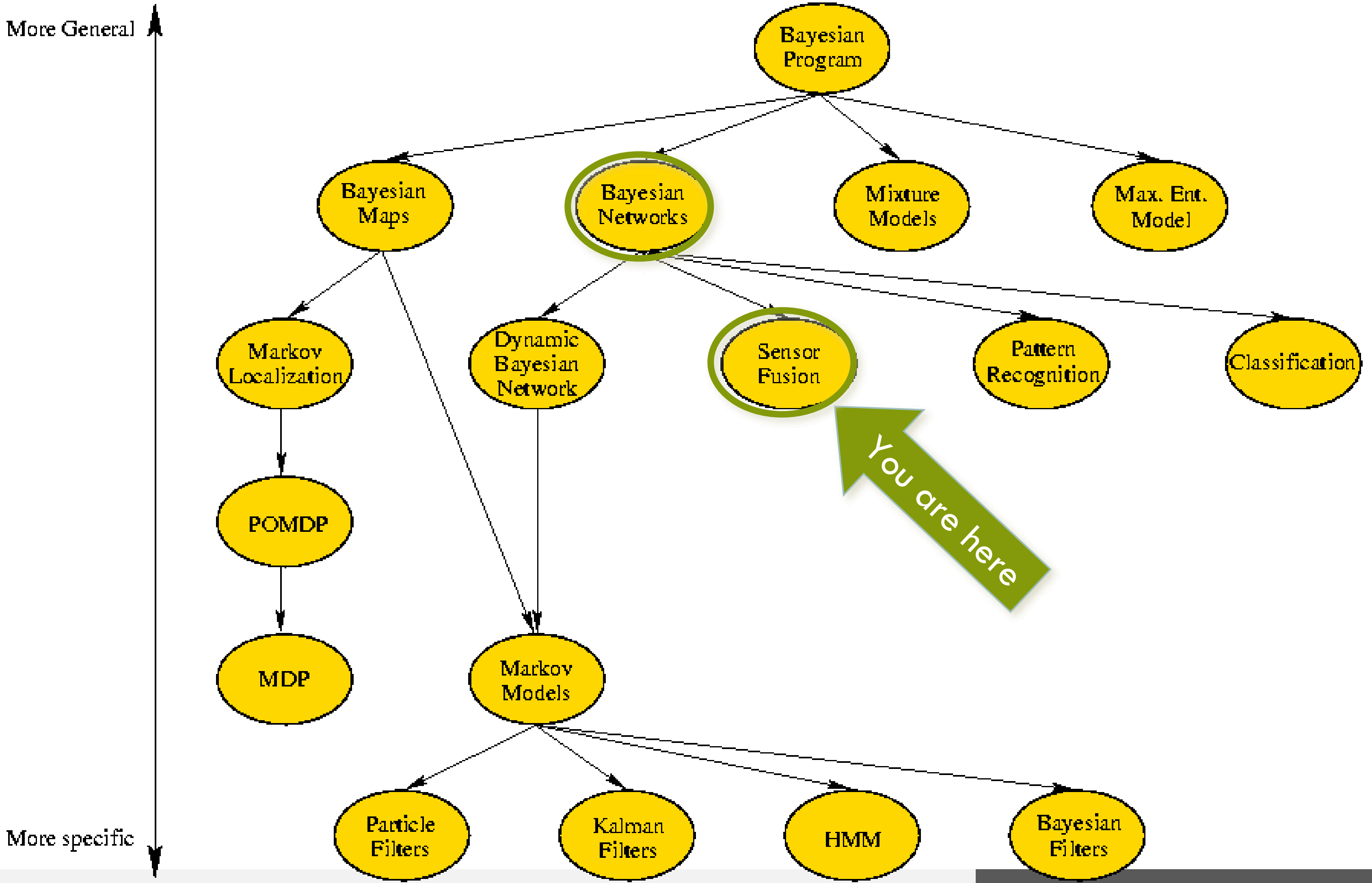
$P(da \mid d) = \text{CndBellShape}(da, d, 10)$

$P(db \mid d) = \text{CndBellShape}(da, d, 15)$



Ejemplo 26: Fusión de 3 sensores

Jerarquía de los programas bayesianos



Naïve Bayes

$$P(D \mid M_1 M_2) = \frac{P(D)P(M_1 M_2|D)}{P(M_1 M_2)}$$

$$\begin{aligned} P(D\ M_1 M_2) &= P(D\ M_1)P(M_2|D\ M_1) \\ &= P(D)P(M_1|D)P(M_2|D\ M_1) \end{aligned}$$

La suposición "ingenua" de independencia consiste en suponer que M_1es condicionalmente de M_2. Dicho de otra manera:

$$P(M_2|D\ M_1) = P(M_2|D)$$

Naïve Bayes

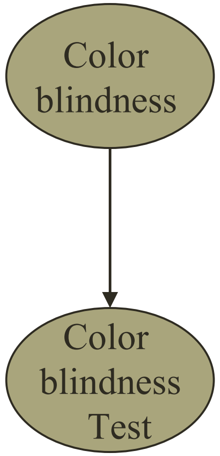
Generalización

$$\begin{aligned}
P(D\ M_1\ M_2 \ldots M_n) &= P(D)P(M_1\ M_2 \ldots M_n \mid D) \\
&= P(D)P(M_1 \mid D)P(M_2\ \ldots M_n \mid D\ M_1) \\
&= P(D)P(M_1 \mid D)P(M_2 \mid D\ M_1)P(M_3\ \ldots M_n \mid D\ M_1) \\
&\vdots \\
&= P(D)P(M_1 \mid D)P(M_2 \mid D\ M_1) \ldots P(M_n \mid D\ M_1\ M_2 \ldots M_n)
\end{aligned}$$

$$P(D \mid M_1\ M_2 \ldots M_n) \propto P(D) \prod_{i=1}^{n} P(M_i \mid D)$$

2.6 Filtros bayesianos

Filtro bayesiano

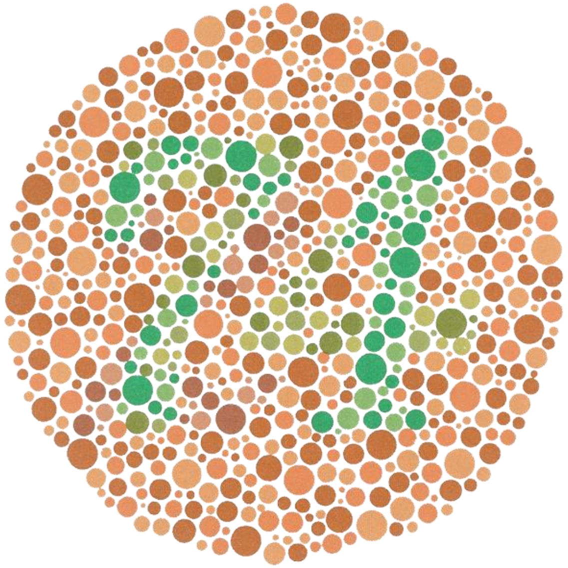


P(CB=f)	P(CB=t)
0.92	0.08

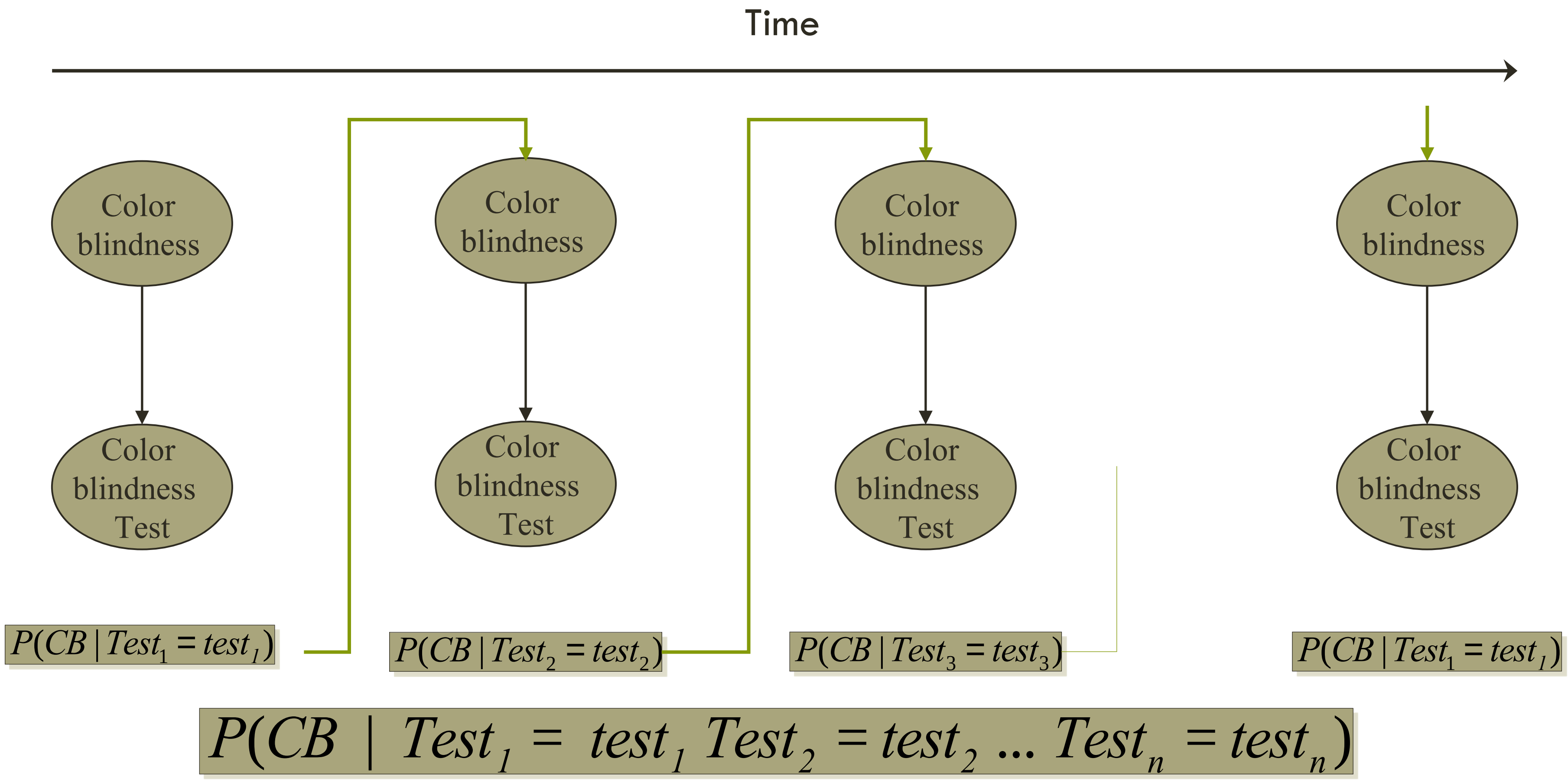
Falso positivo

CB	P(T=f | CB)	P(T=t | CB)
f	0.8	0.2
t	0.1	0.9

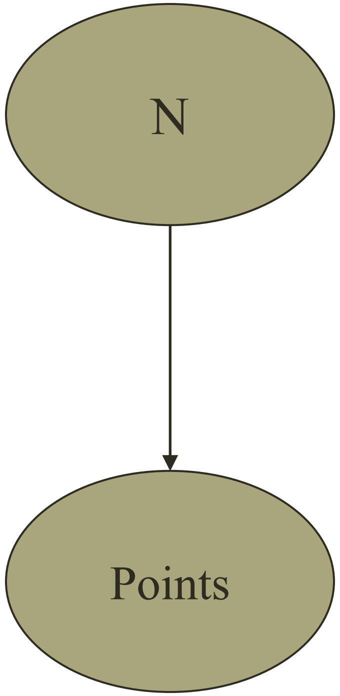
Falso negativo



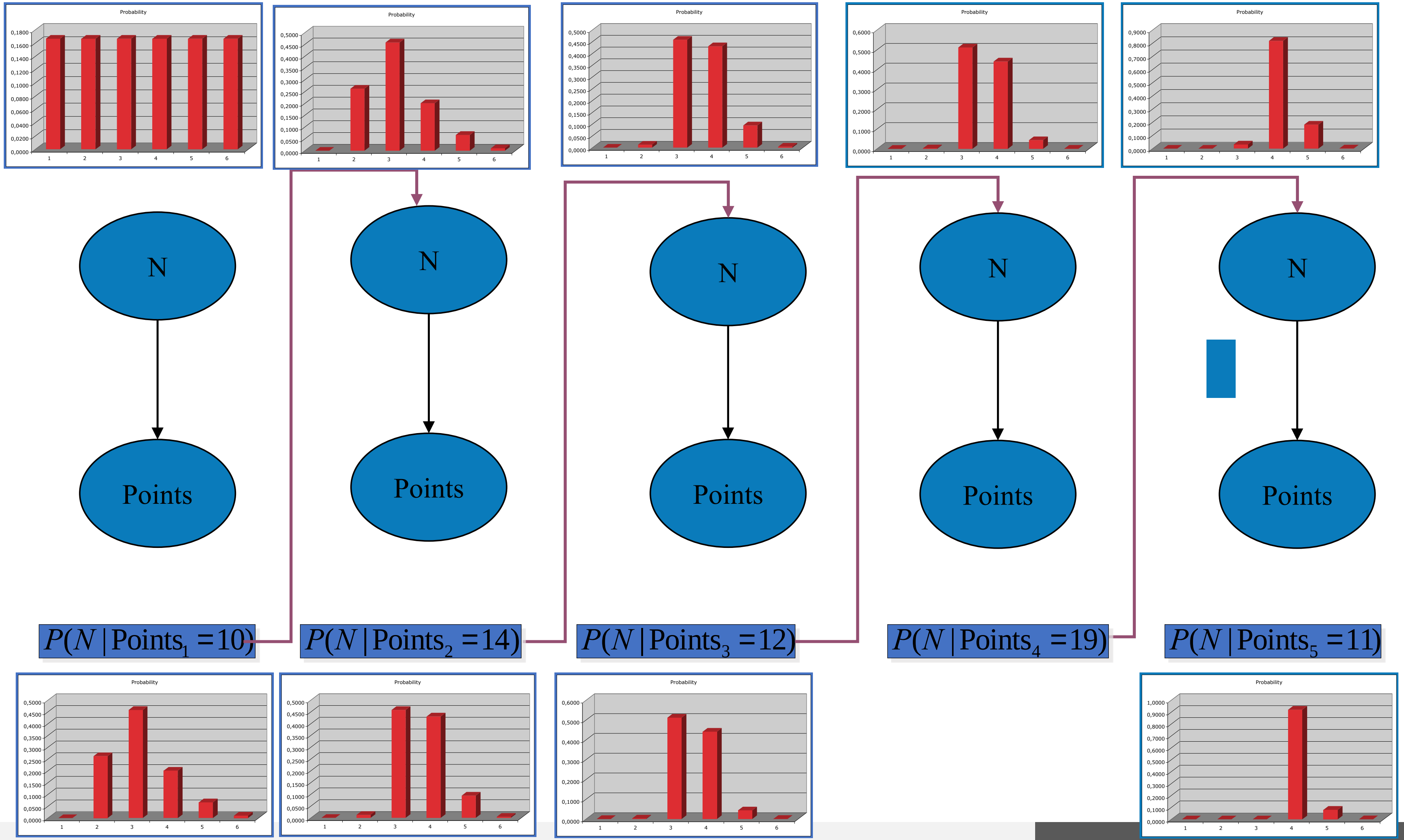
Filtro bayesiano



Filtro bayesiano



Filtro bayesiano



Filtro bayesiano

Considera que tu compañera tiene una caja con m dados; en secreto coge $n \leq m$ dados y los lanza sobre la mesa. Te hace saber que la suma de todos los puntos es s. ¿Cuál es el número más probable n de dados.

Se sabe que cuando n es particularmente "grande", tenemos que la distribución del número de puntos puede ser aproximada por una gaussiana con

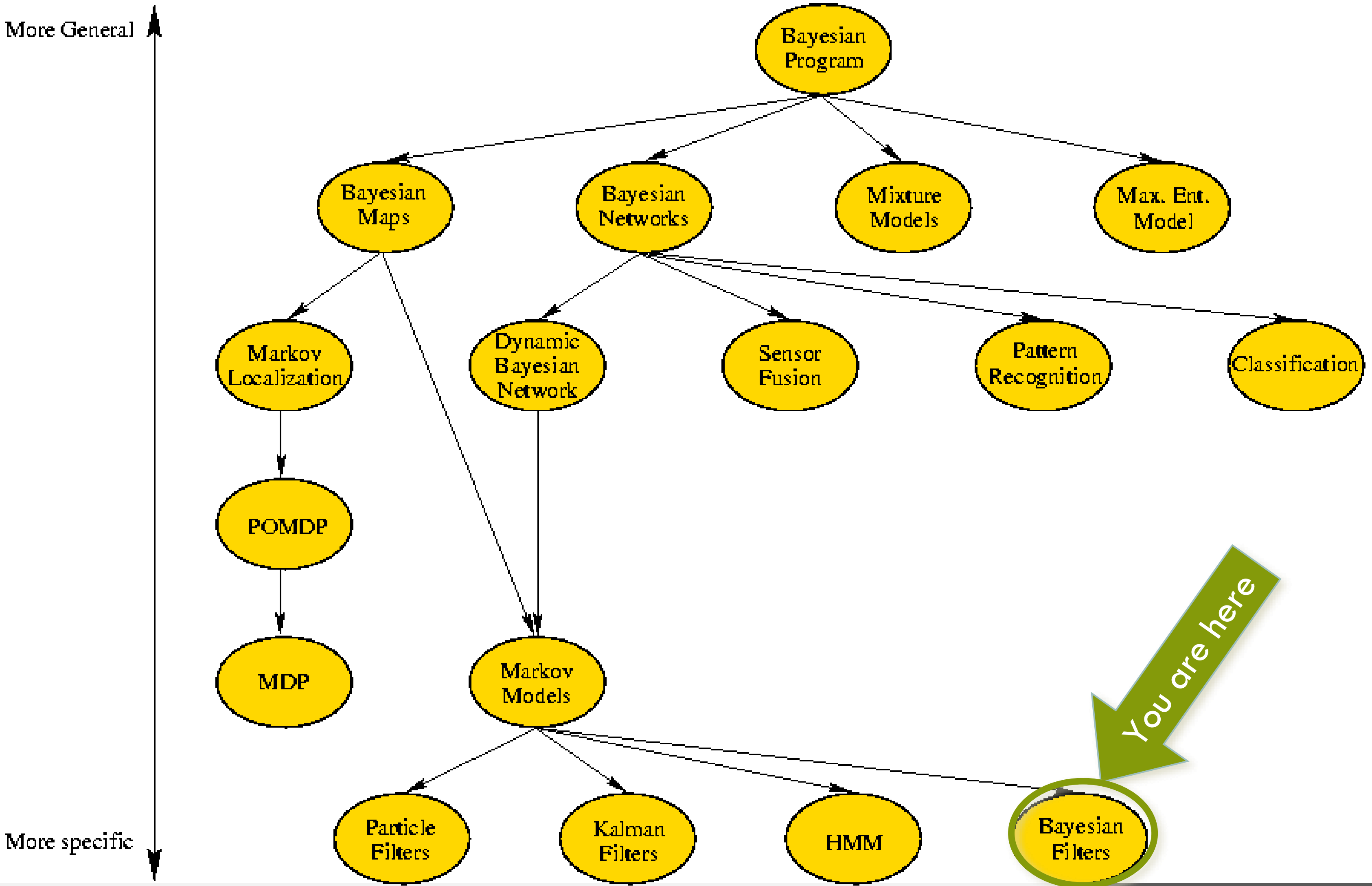
$$m = \frac{7n}{2}$$

y

$$s = \sqrt[2]{\frac{35n}{12}}$$

Ejemplo 27: Filtro bayesiano

Jerarquía de los programas bayesianos



Filtro bayesiano

[Akinator. (Noviembre 2017). http://es.akinator.com/](http://es.akinator.com/)

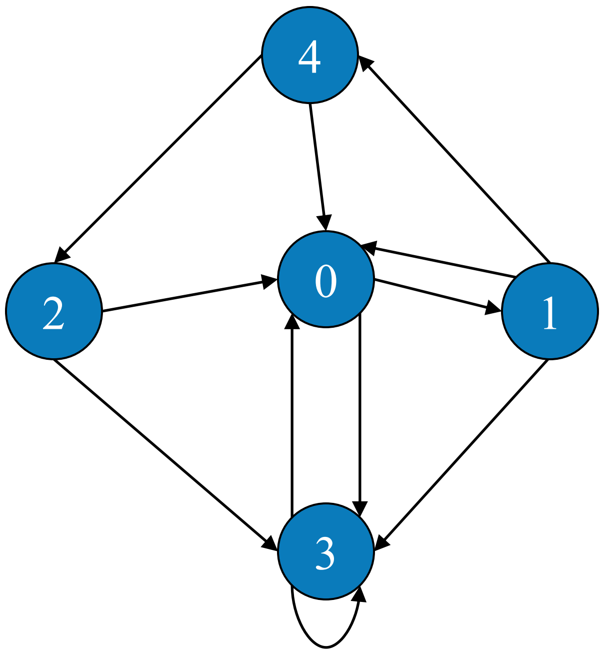
2.7 Modelos ocultos de Markov

Modelos ocultos de Markov (HMM)

$$P(S_t\ S_{t-1}\ O_t) = P(S_{t-1})\ P(S_t \mid S_{t-1})\ P(Ot \mid S_t)$$

Modelos ocultos de Markov (HMM)

Grafo de transición



$$P(S_t \mid S_{t-1})$$

	$S_t=0$	$S_t=1$	$S_t=2$	$S_t=3$	$S_t=4$
$S_{t-1}=0$	0.0	0.2	0.0	0.8	0.0
$S_{t-1}=1$	0.5	0.0	0.0	0.3	0.2
$S_{t-1}=2$	0.5	0.0	0.0	0.5	0.0
$S_{t-1}=3$	0.4	0.0	0.0	0.6	0.0
$S_{t-1}=4$	0.3	0.0	0.7	0.0	0.0

Modelos ocultos de Markov (HMM)

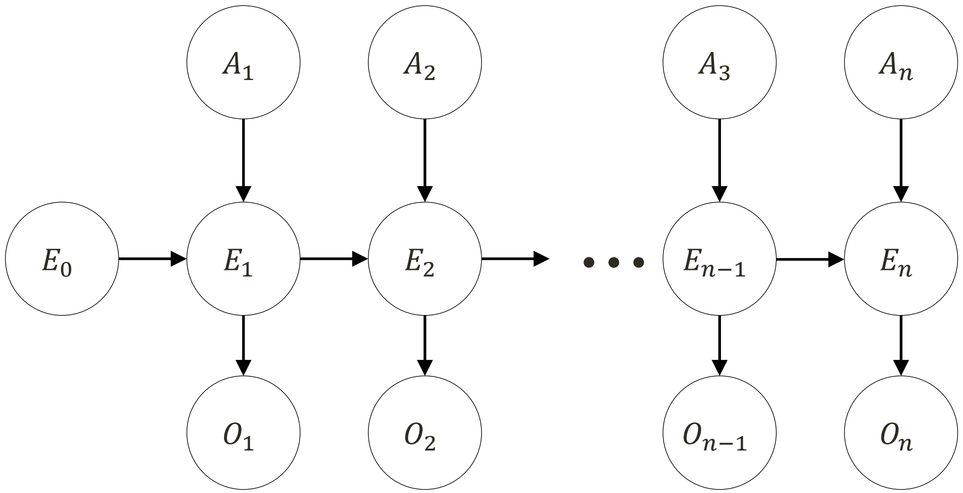
Observaciones

$$P(Ot \mid S_t)$$

	O=0	O=1	O=2	O=3	O=4
S_t=0	0.3	0.3	0.2	0.1	0.1
S_t=1	0.2	0.2	0.2	0.3	0.1
S_t=2	0.6	0.1	0.0	0.3	0.0
S_t=3	0.2	0.2	0.3	0.0	0.3
S_t=4	0.8	0.1	0.1	0.0	0.0

Modelos ocultos de Markov (HMM)

El modelo

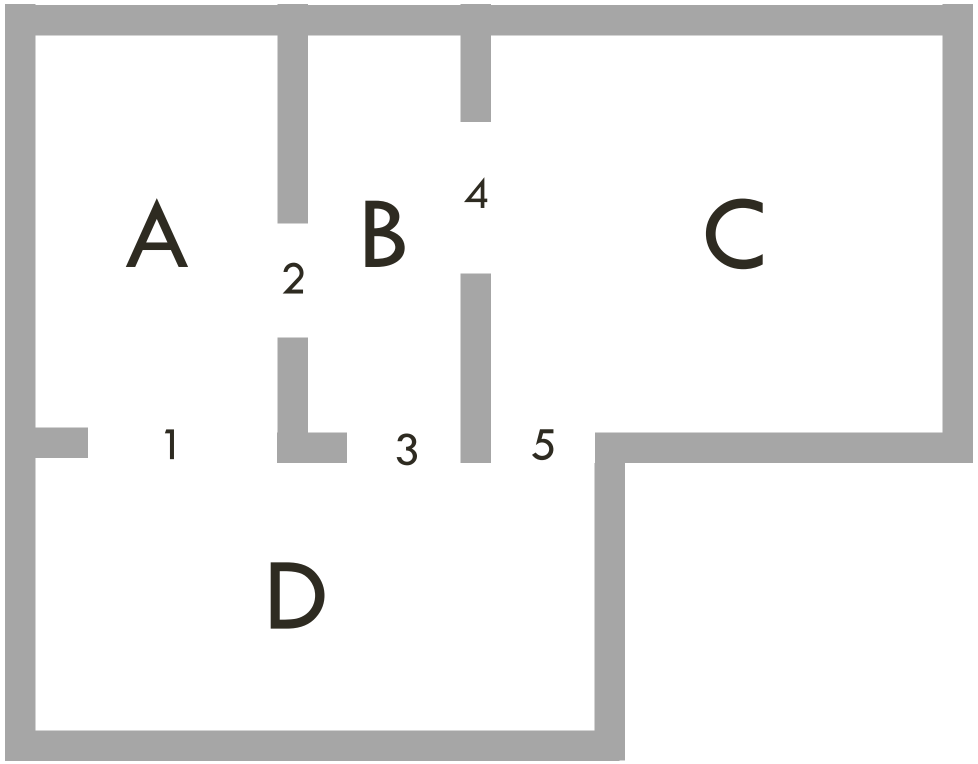


Modelos ocultos de Markov (HMM)

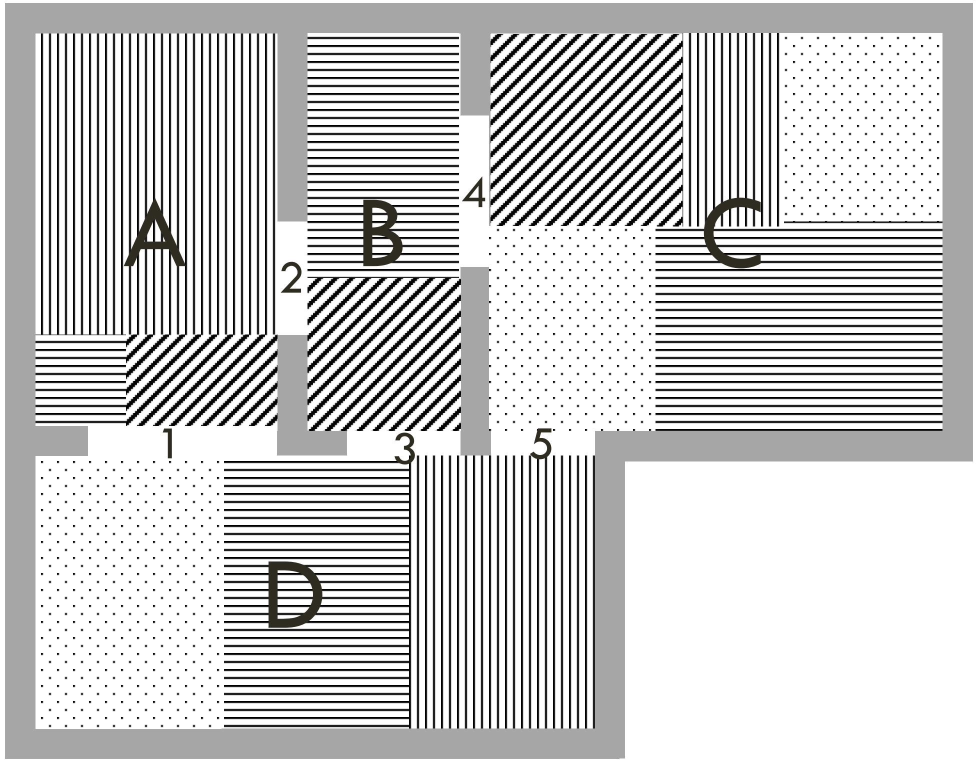
Especificación del programa

$$
Filter(t,o) = \left\{ \begin{array}{l}
\text{Description} \left\{ \begin{array}{l}
\text{Specification} \left\{ \begin{array}{l}
\text{Relevant Variables:} \\
State_p, State, Action, Obs \\
\text{Decomposition:} \\
P(State_p \ Action \ State \ Obs \mid \pi) = \\
\quad P(State_p \mid \pi) P(Action \mid \pi) \\
\quad P(State \mid State_p \ Action \ \pi) P(Obs \mid State \ \pi) \\
\text{Parametric Forms:} \\
\ \ \text{Initial PF} \\
\quad P_0(State_p \mid \pi) = \gamma_0 \\
\ \ \text{Prior PF} \\
\quad P(State_p \mid \pi) = \left\{ \begin{array}{ll} rename(State_p, Filter(t-1, o)) & \text{if } t \neq 1 \\ P_0(State_p \mid \pi) & \text{otherwise} \end{array} \right. \\
\ \ \text{Transitive PF} \\
\quad P(Action \mid \pi) = \gamma_A \\
\quad P(State \mid State_p \ Action \ \pi) = \gamma_S \\
\quad P(Obs \mid State \ \pi) \gamma_{Obs}
\end{array} \right. \\
\text{Identification:}
\end{array} \right. \\
\text{Question:} \\
\ P(State \mid Obs = o[t] \ \pi)
\end{array} \right.
$$

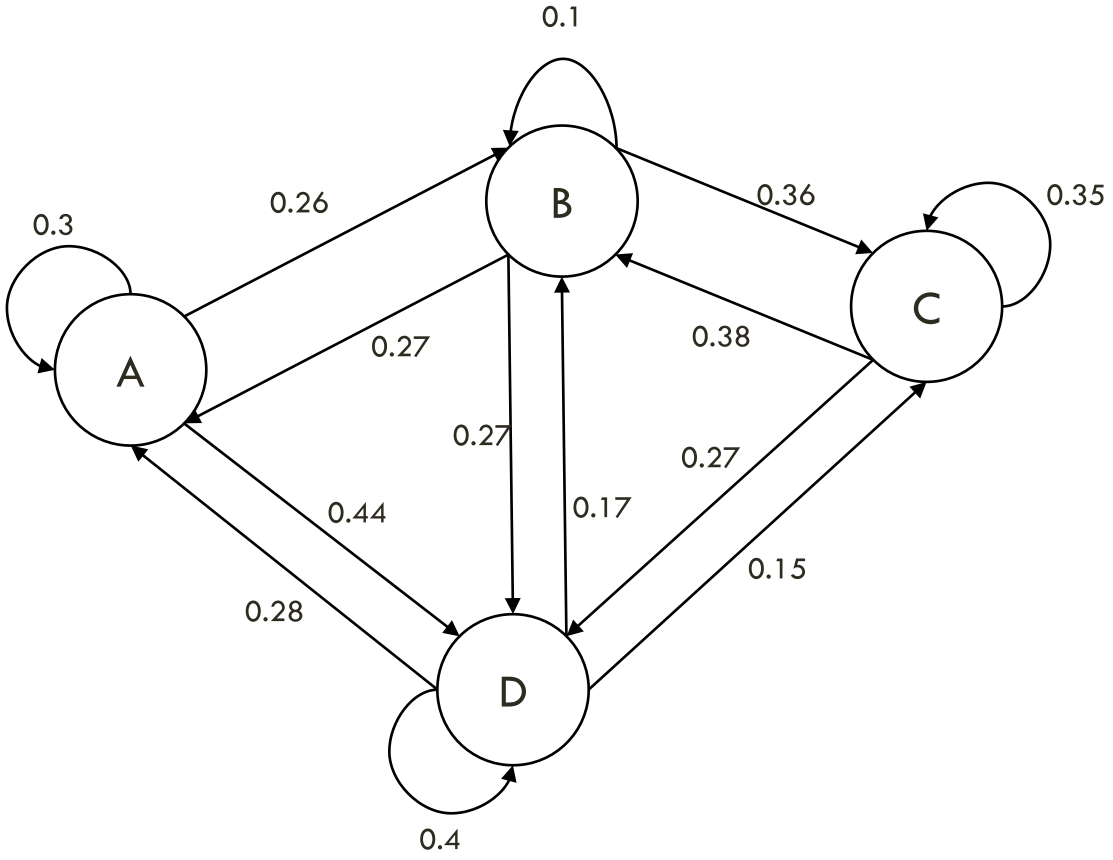
Modelos ocultos de Markov (HMM)



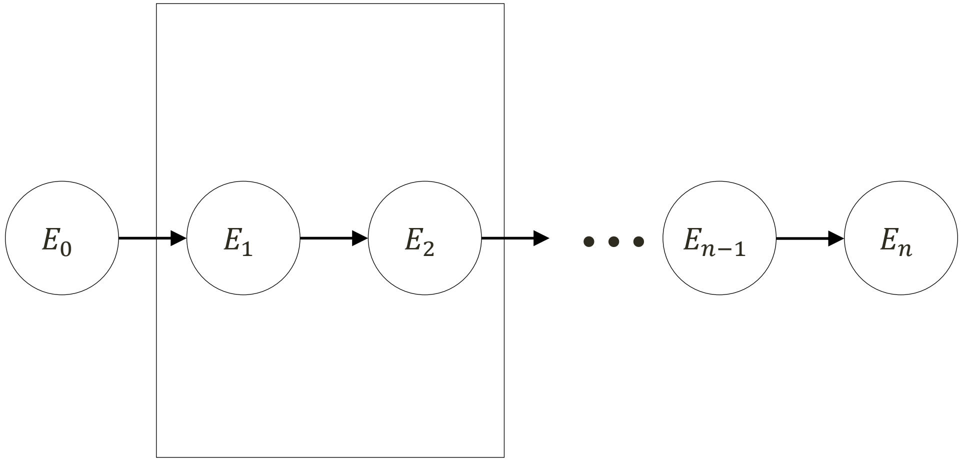
Modelos ocultos de Markov (HMM)



Modelos ocultos de Markov (HMM)



Modelos ocultos de Markov (HMM)



Modelos ocultos de Markov (HMM)

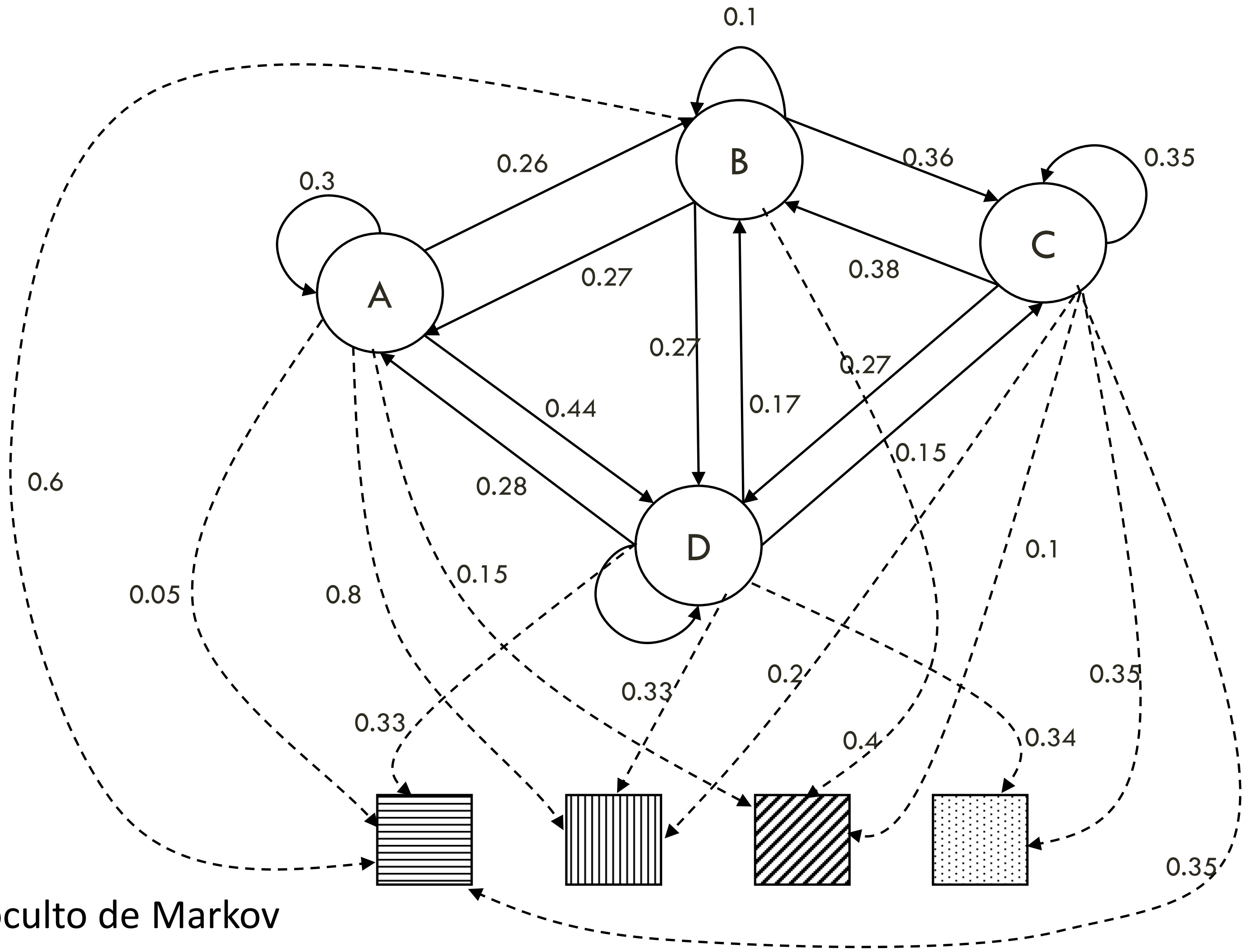
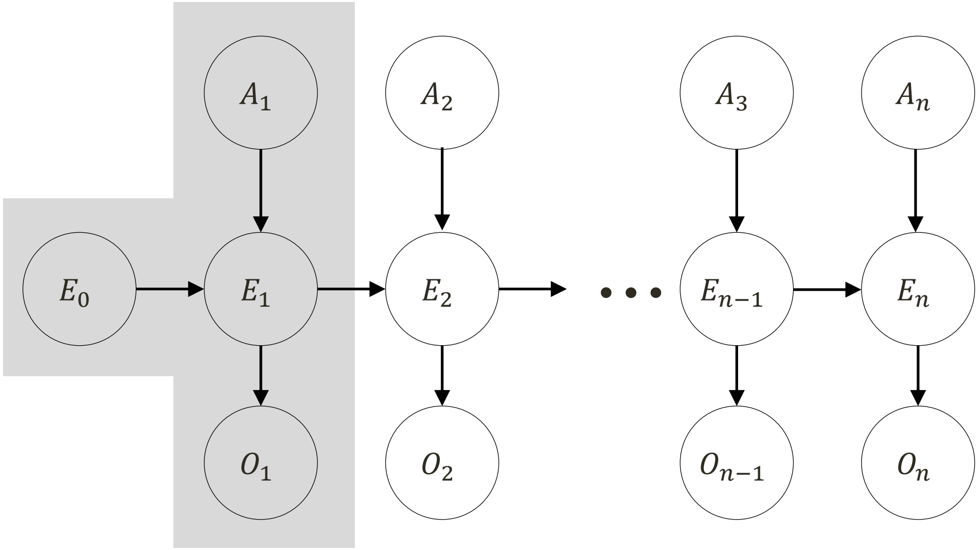
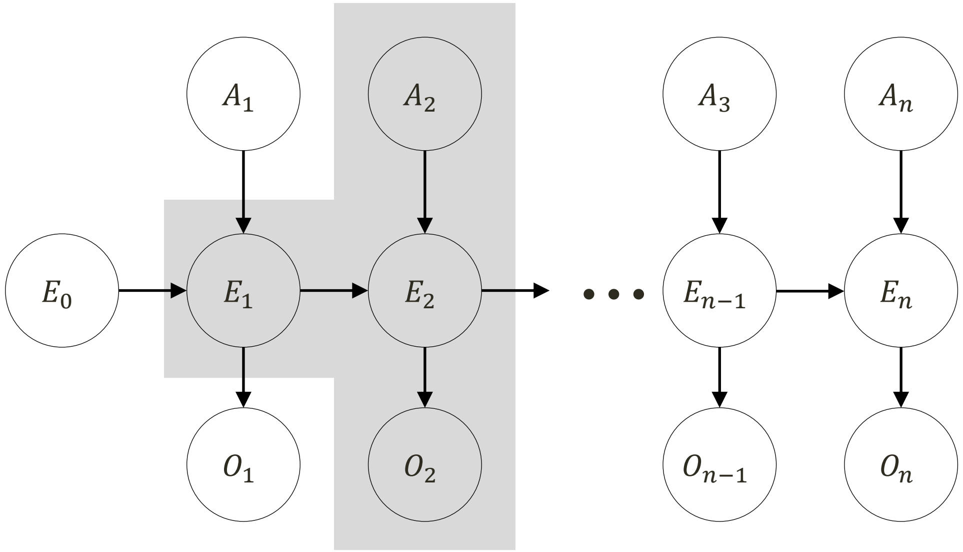


Diagrama del Modelo oculto de Markov

Modelos ocultos de Markov (HMM)

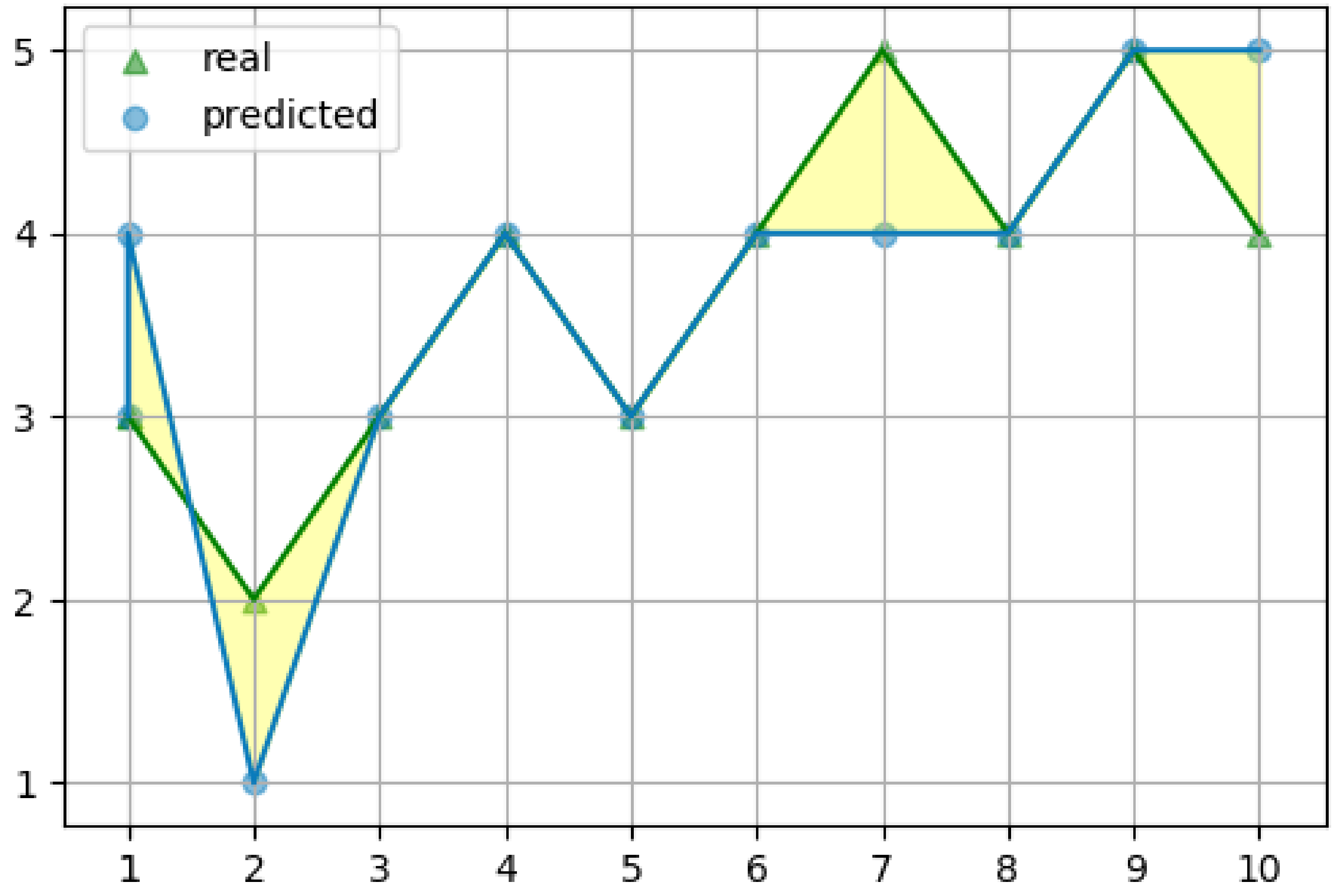


Modelos ocultos de Markov (HMM)



Ejemplo 28: HMM

Modelos ocultos de Markov (HMM)



Jerarquía de los programas bayesianos