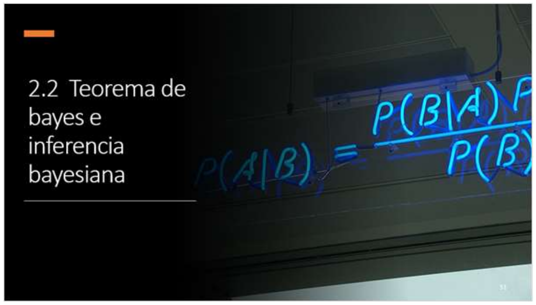
Redes Neuronales (y Bayesianas)
LDS1081

Juan Manuel Ahuactzin Larios
juan.ahuactzin@udlap.mx

Contenido: Da clic en la sección que quieres consultar.



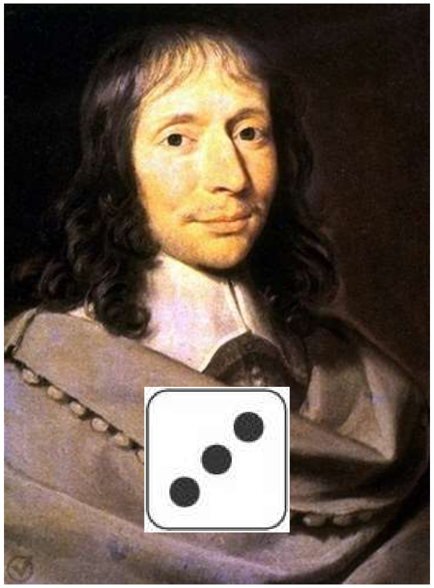
2.1 Incompetud vs Incertidumbre

Juan Manuel Ahuactzin Larios

juan.ahuactzin@udlap.mx

Experimentos aleatorios

1654

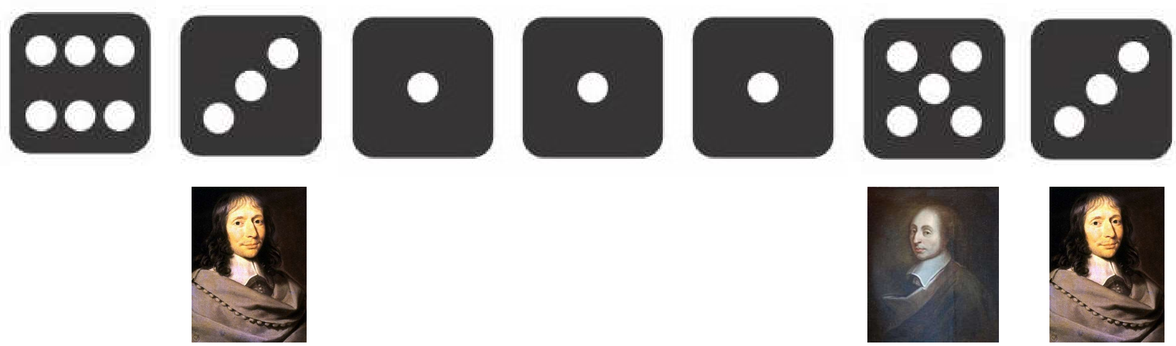


Antoine de Gombard

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

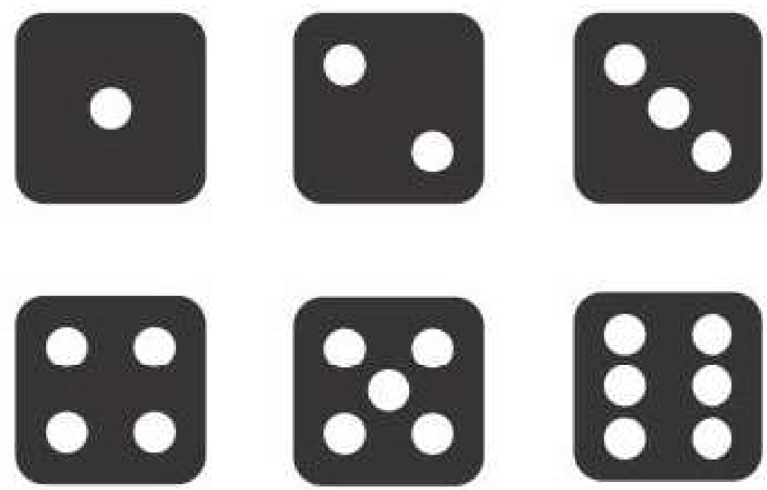
Experimentos aleatorios

Da ejemplos de eventos deterministas y aleatorios

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

Espacio muestral

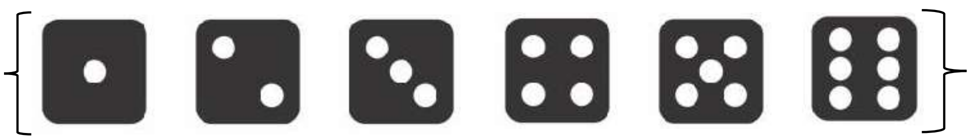
En un evento aleatorio no podemos saber el resultado pero si el conjunto de todos los resultados posibles.



Espacio muestral

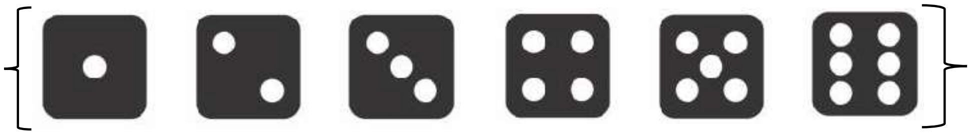
Espacio muestral es el **conjunto** de todos los resultados posibles del experimento y se denota por **Ω** (Omega mayúscula).

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.

Sorteo
MAYOR
3714
0 0 0 0 0
SERIE 0
Carla Elena
Lotería Nacional
para la Asistencia Pública
Un total de
54
MILLONES
de pesos en premios*
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

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

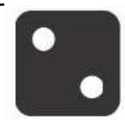
Espacio muestral

¿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.

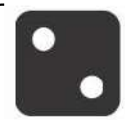
$\Omega_A = \{$  $\}$

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

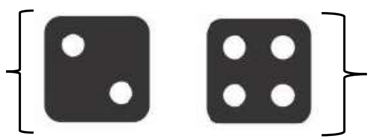
$\Omega_C=$ {  } Obtener un número divisible entre 2 y 3

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)

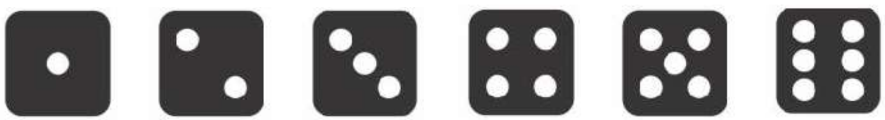
$\Omega_A = \{ \quad \}$ Obtener un número divisible entre 2 y 5

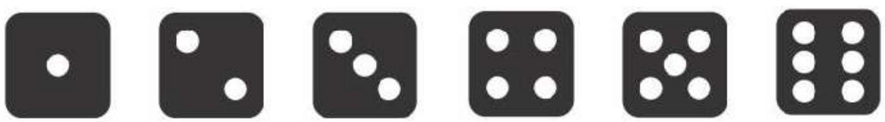
$\Omega_B = \{ \quad \}$ Obtener un número par menor que 1

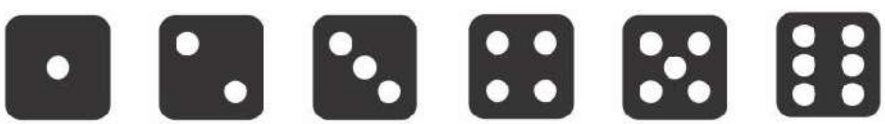
$\Omega_C = \{ \quad \}$ Obtener un número mayor a 6

Espacio muestral

Evento seguro

$\Omega_A =$  Obtener un número menor que 7

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

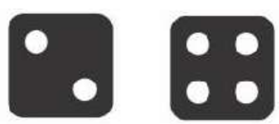
$\Omega_B =$  Obtener un número entre 1 y 6

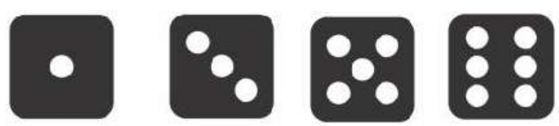
Espacio muestral

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

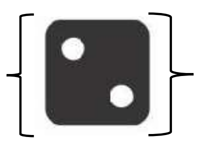
https://docs.google.com/forms/d/e/1FAIpQLScIaxcMXSiO6jy96CZrnpSkIRkHAFuOQSFjlvyZWUm6Hs9yhw/viewform

Operaciones con conjuntos

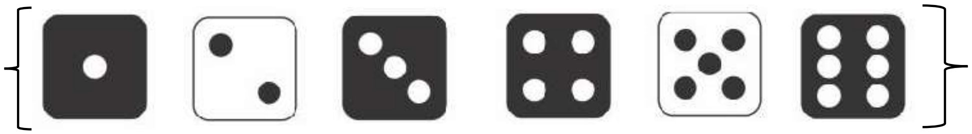
Para calcular la **probabilidad** de ocurrencia de cada evento es indispensable identificar cada uno de ellos.

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.

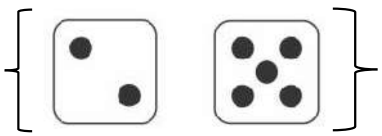
Operaciones con conjuntos

$\Omega =$  Es el conjunto de todos los resultados de **lanzar un dado**.

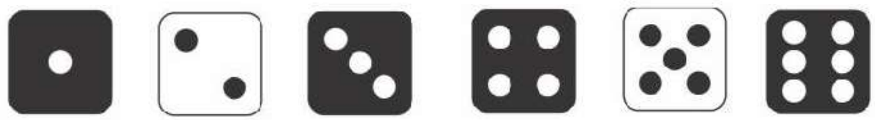
Mi número es primo

$\Omega_A =$ 

Mi número es blanco

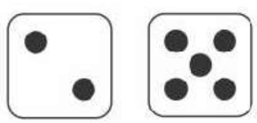
$\Omega_{A_1} =$ 

Operaciones con conjuntos

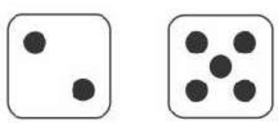
$\Omega = \{$  $\}$

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

Mi número es blanco

$\Omega_B = \{$  $\}$

Mi número es primo

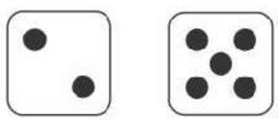
$\Omega_{B_1} = \{$  $\}$

Operaciones con conjuntos

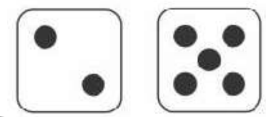
Sabiendo que mi número es primo que proporción es blanco.

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

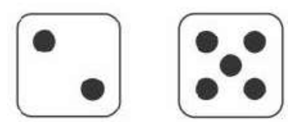
$\frac{2}{3}$

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

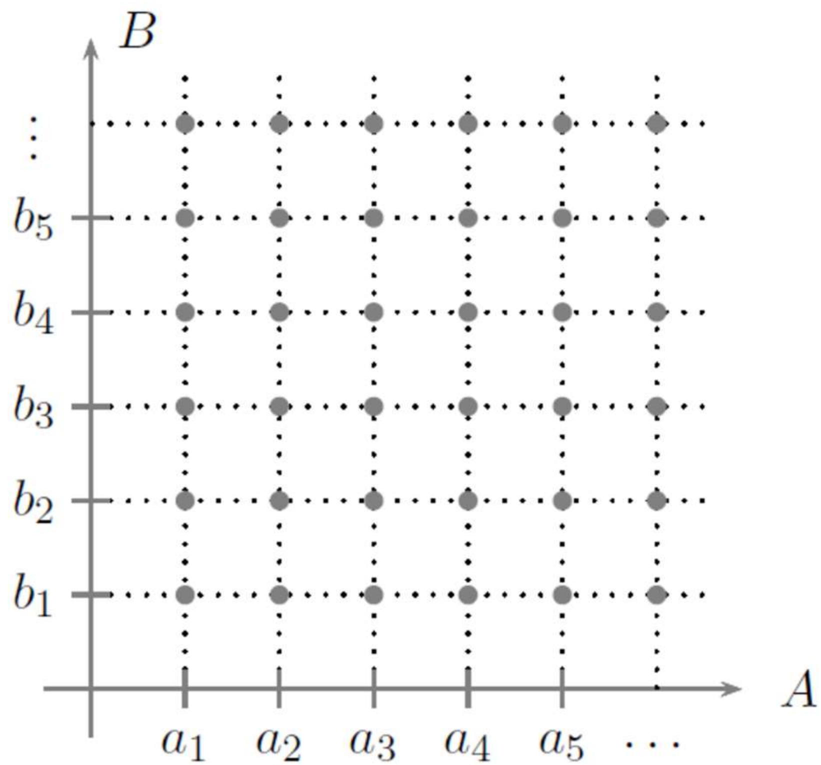
El producto cartesiano de dos conjuntos A y B se define como el conjunto de todas las parejas ordenadas de (a,b) con a perteneciente a A y b perteneciente a B y se denota por " $\times$ "

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

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

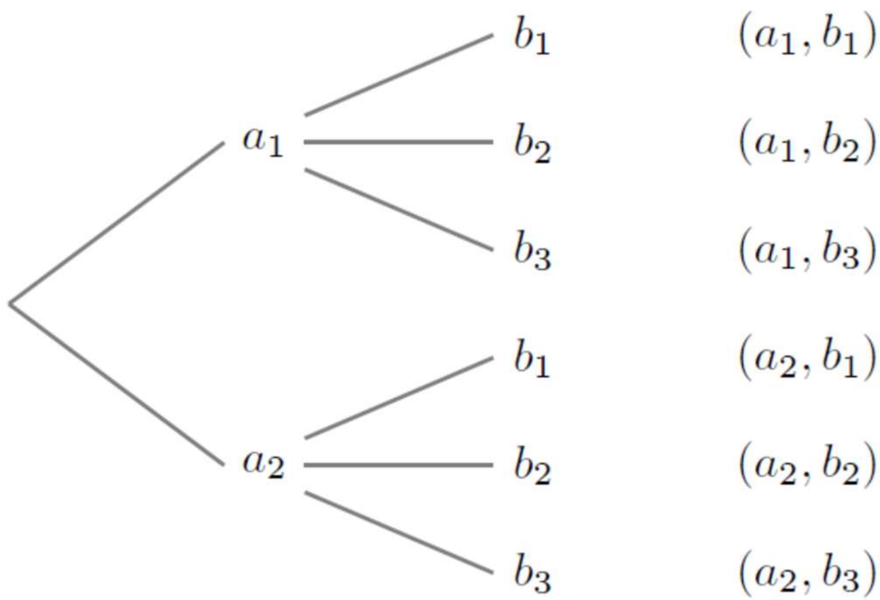
Operaciones con conjuntos

Representación del Producto cartesiano



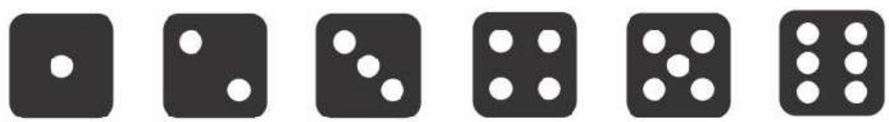
Operaciones con conjuntos

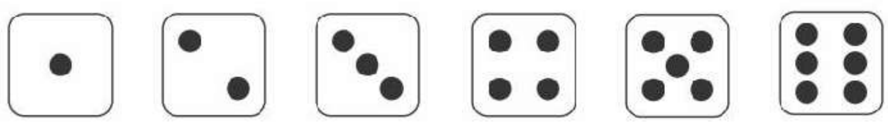
Representación del Producto cartesiano



Operaciones con conjuntos

Representación del Producto cartesiano

$\Omega_N = \{$  $\}$

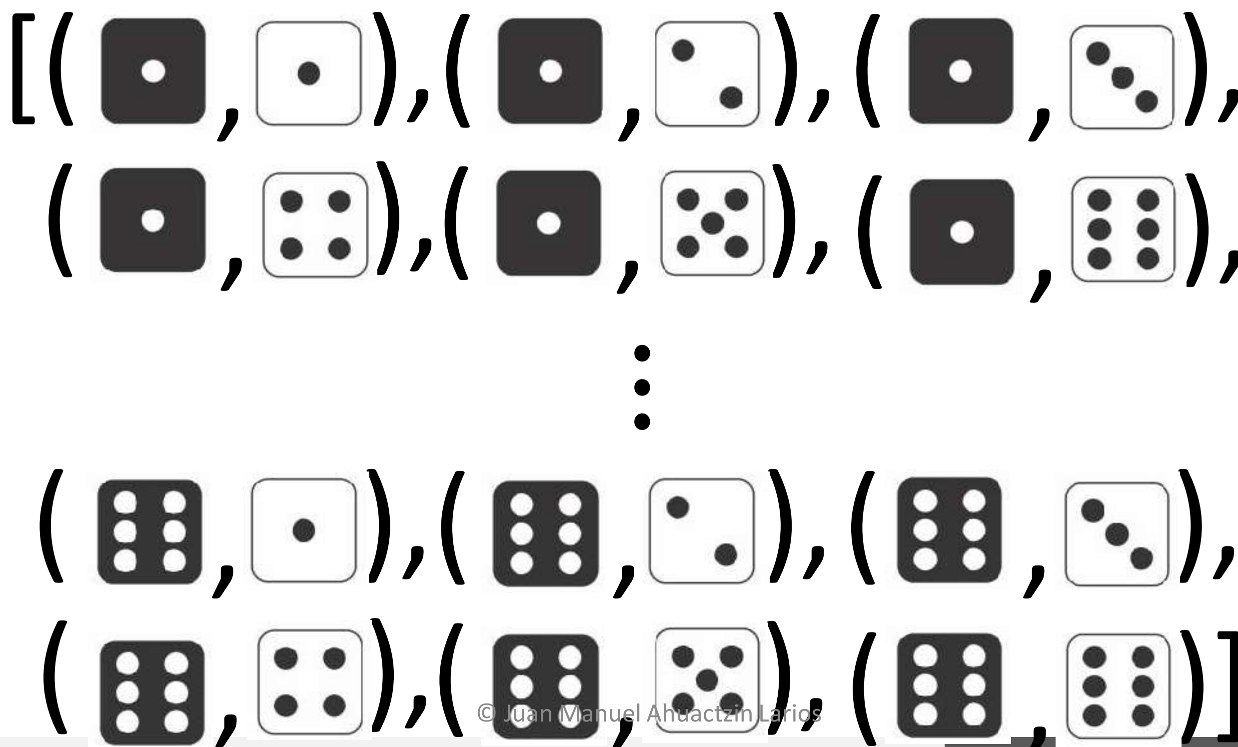
$\Omega_B = \{$  $\}$

$$\Omega_N \times \Omega_B$$

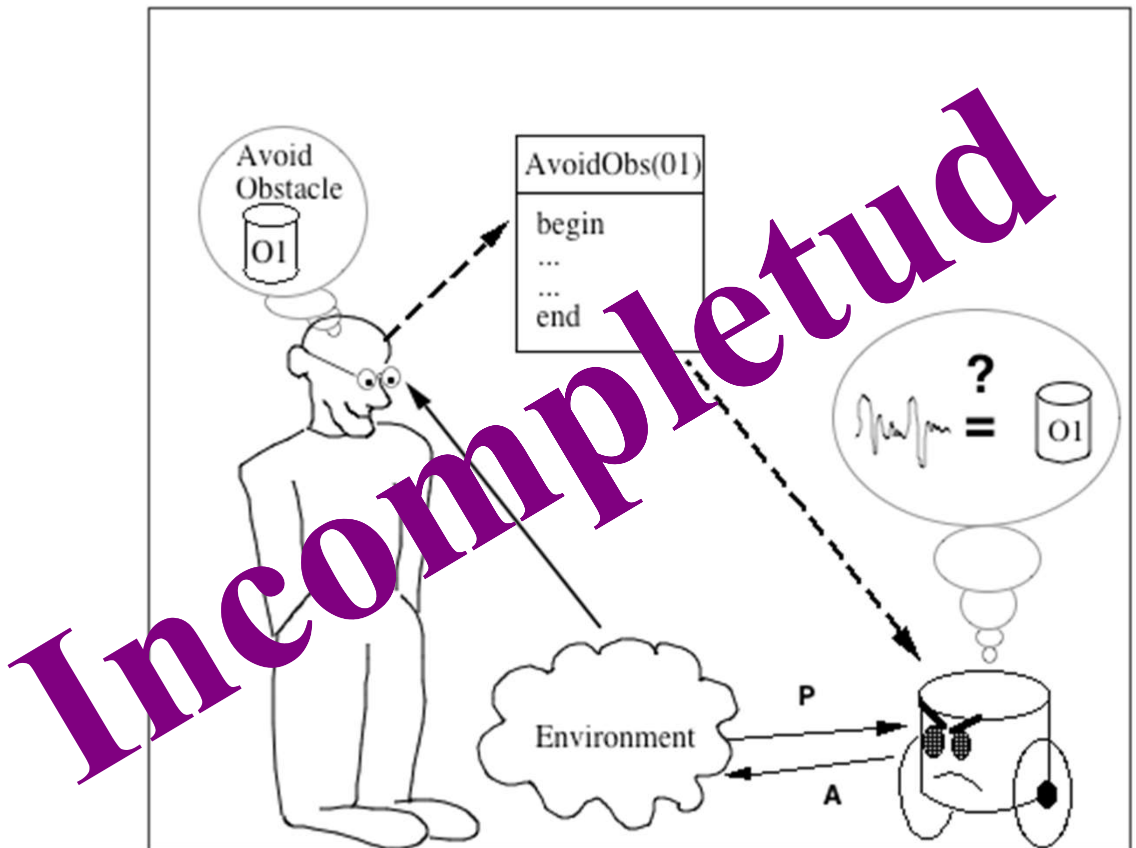
Operaciones con conjuntos

Representación del Producto cartesiano

$\Omega_N \times \Omega_B =$

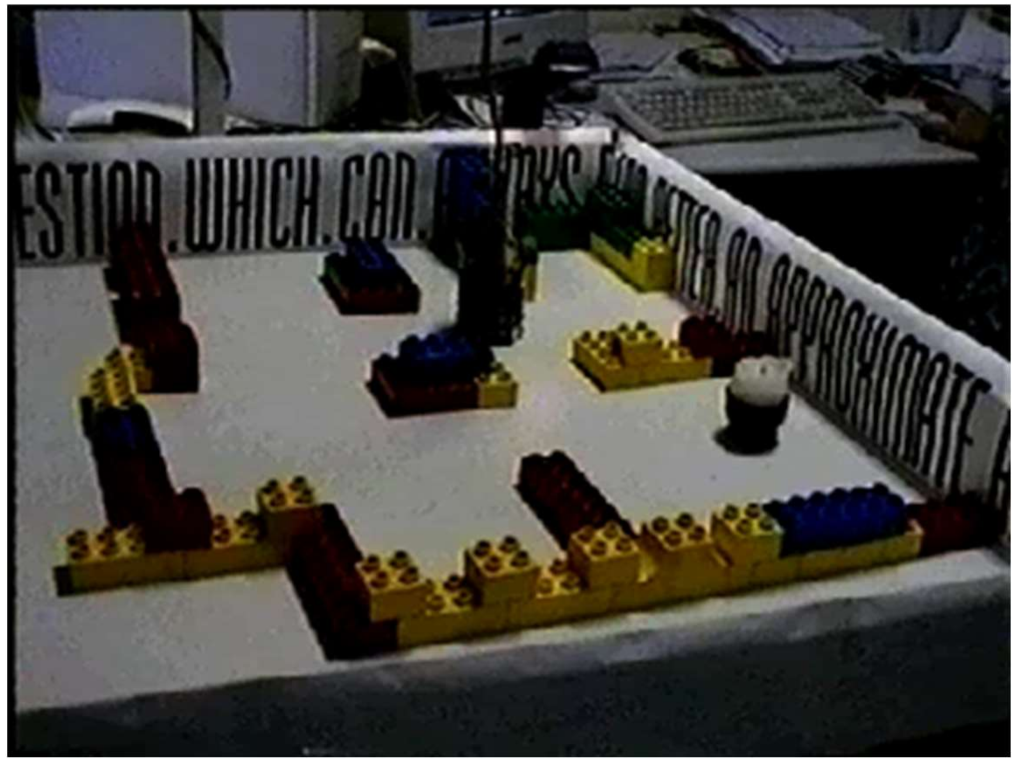


Paradigma lógico



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

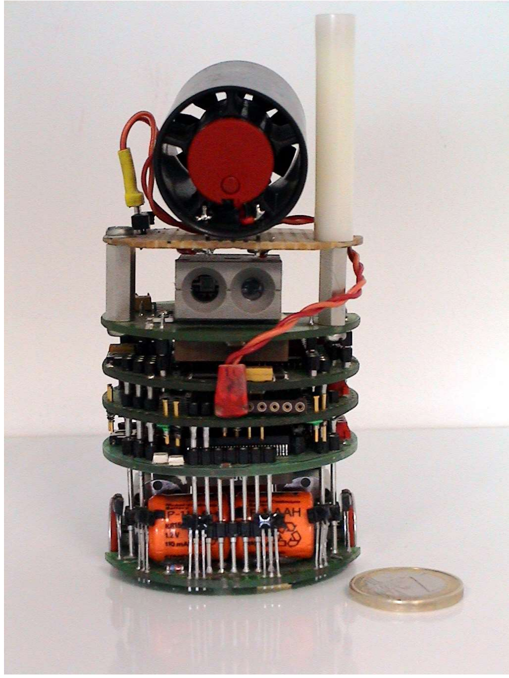
Photo Alain Herzog



Aprendiendo comportamientos reactivos

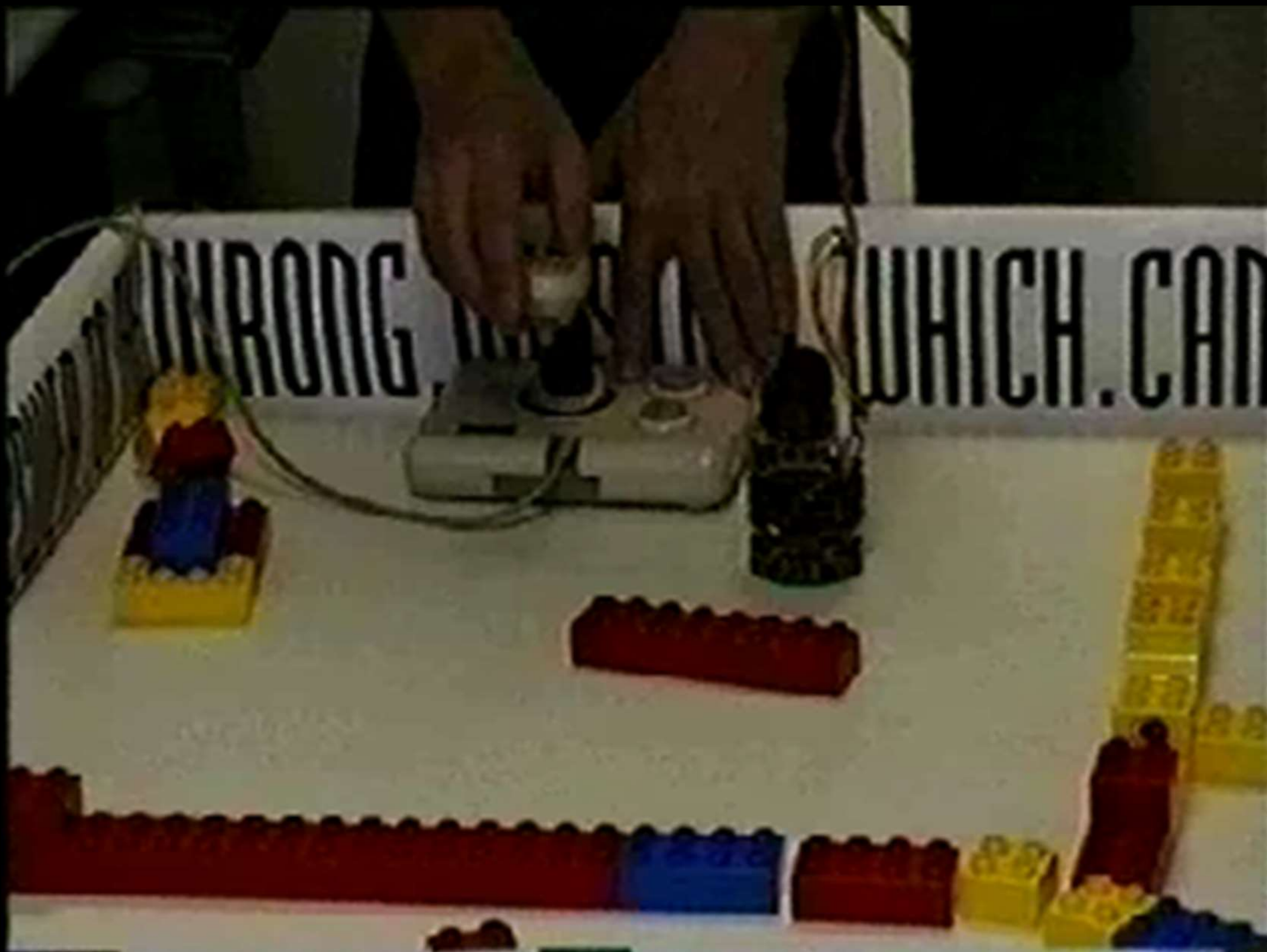
Lebeltel, O., Bessière, P., Diard, J. & Mazer, E. (2004) Bayesian Robot Programming; *Autonomous Robots*, Vol. 16, p. 49-79

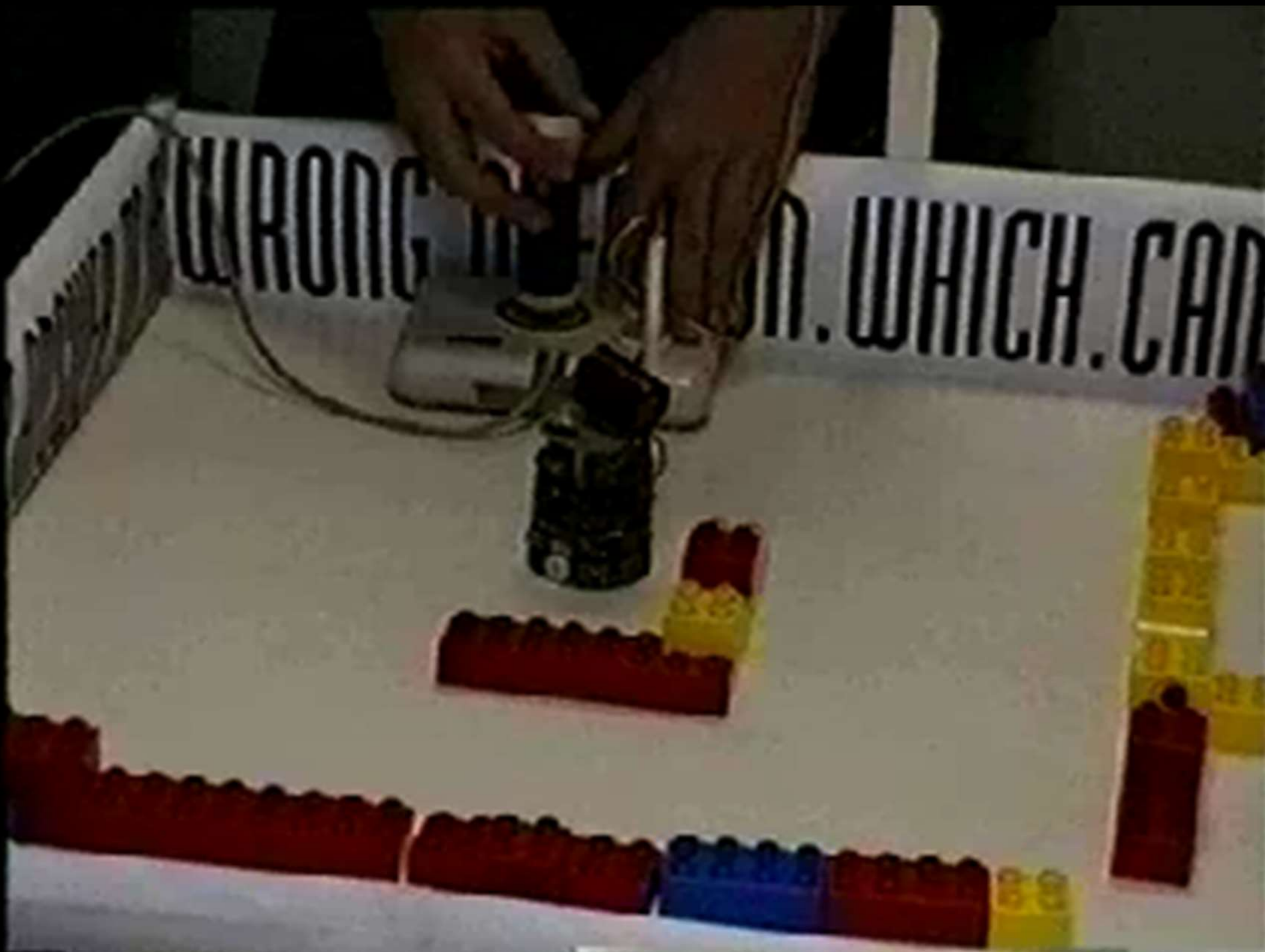
Lebeltel, O. (1999) *Programmation bayésienne des robots*; Thèse INPG

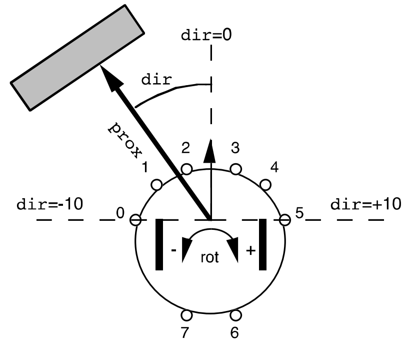


Khepera

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




WRONG
WHICH.



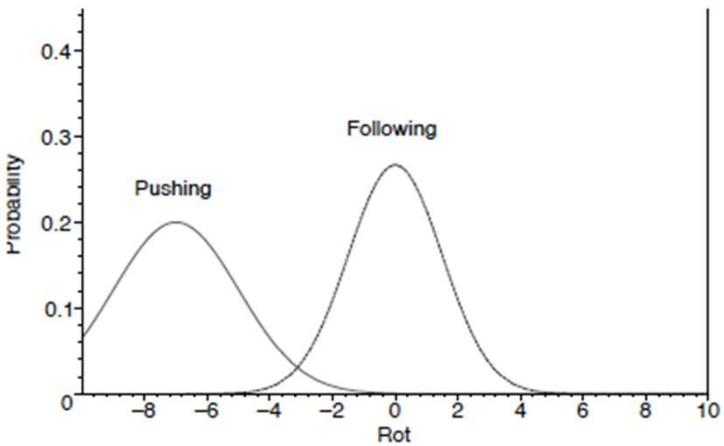


Figure 4.4: $\mathbf{P}(Rot \mid [Dir = -10] \wedge [Prox = 13])$

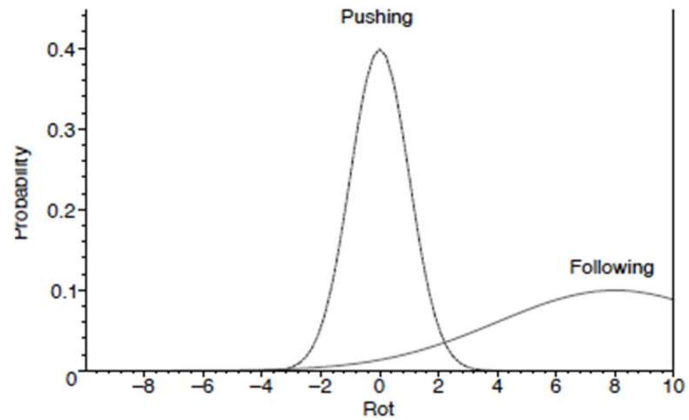


Figure 4.5: $\mathbf{P}(Rot \mid [Dir = 0] \wedge [Prox = 15])$

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:

$$P(b \mid a) = 1$$

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

  - Modus Tollens:

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

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

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

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, $P(a \mid b) = \frac{3}{10}$
por que **3**, **6**, **9**, **12**, **15**, **18**, **21**, **24**, **27**, **30**, … } 3 de los 10 números

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

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

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

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, $P(b) = \frac{10}{30} = \frac{1}{3}$

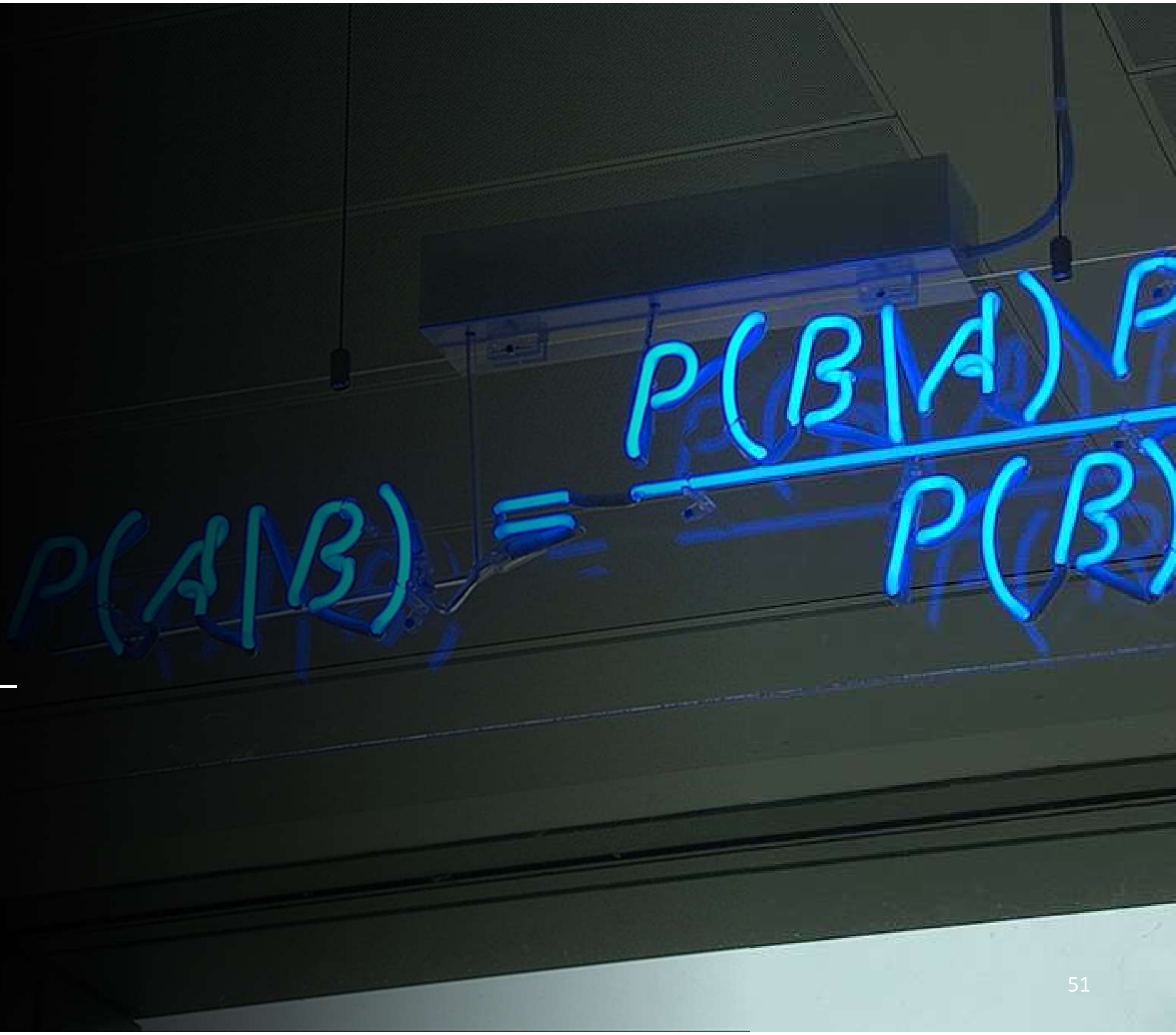
1, 2, **3**, 4, 5, **6**, 7, 8, **9**, 10, 11, **12**, 13, 14, **15**, 16, 17, **18**, 19, 20, **21**, 22, 23, **24**, 25, 26, **27**, 28, 29, **30**, …

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: $P(b \mid a) = 1 \Rightarrow P(b \mid \neg a) \leq P(b) = \frac{7}{27} \leq \frac{1}{3} = 0.259 \leq 0.33333$

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

$$P(XY)$$

$$P(X)$$

$$P(Y)$$

$$P(X \mid Y)$$

$$P(Y \mid X)$$

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

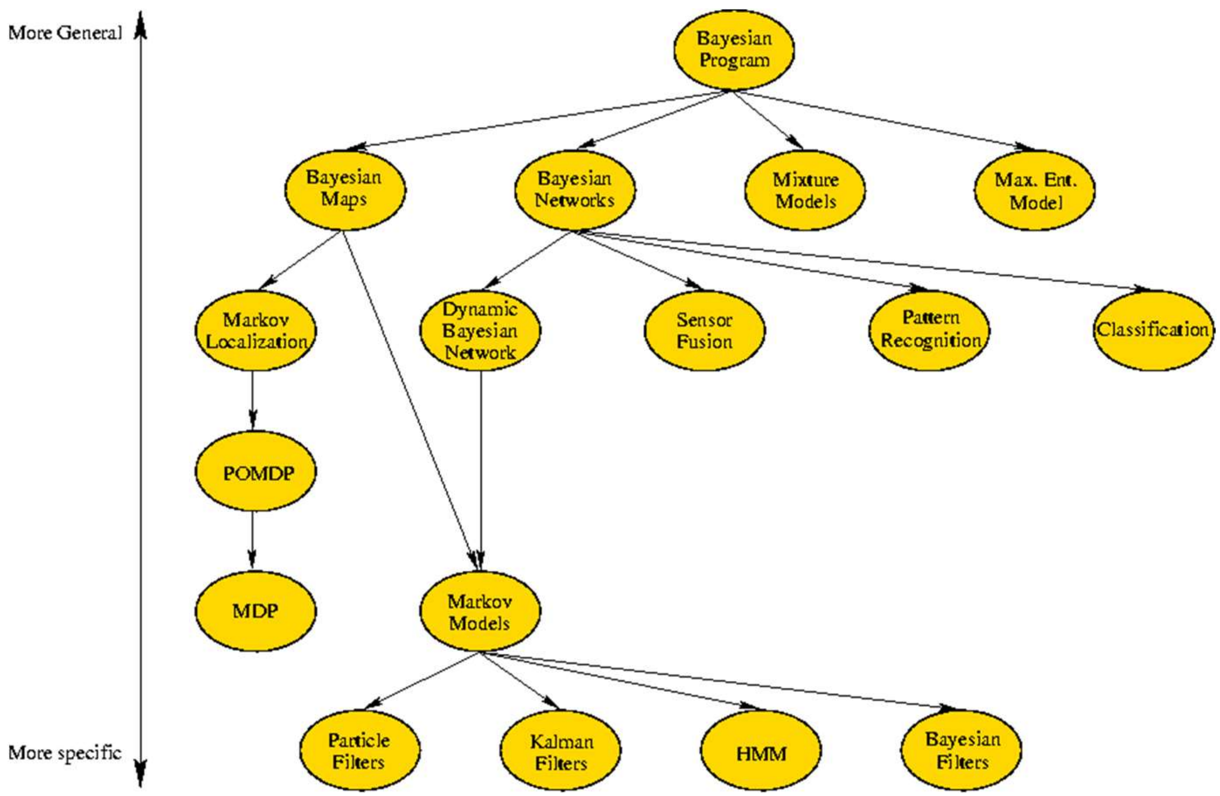
Trabajemos con

$$\mathrm{P(Color\ Forma) = P(Forma)P(Color|\ Forma)}$$

$$P(Forma = Luna) = 0.61$$
$$P(Forma = Estrella) = 0.39$$

Construct $P(Colors \mid Forma)$

Jerarquía de los programas bayesianos



Ejemplo 14: Inferencia con el Teorema de Bayes

2.3 Redes bayesianas

Bayesian program = Description + Question

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

Ejercicio 06: Dibujar las redes bayesianas

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

$BiasedDice() =$

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:

$P(Die|Points)$

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

$Spam$

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

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

$$P(Spam \wedge W_n)$$

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

Postulado de la conjunción

$$
\begin{aligned}
P(X \wedge Y) &= P(X) \times P(Y \mid X) \\
&= P(Y) \times P(X \mid Y)
\end{aligned}
$$

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

$$
\begin{aligned}
P([Spam = true] \wedge [W_n = true]) &= P([Spam = true]) \times P([W_n = true] \mid [Spam = true]) \\
&= 0.75 \times 0.0004 \\
&= 0.0003 \\
&= P([W_n = true]) \times P([Spam = true] \mid [W_n = true])
\end{aligned}
$$

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)

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

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

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

Distribución conjunta y preguntas (3)

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

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

Decompocición

$$
\begin{aligned}
& P(Spam \wedge W_0 \wedge \ \dots \ \wedge W_{N-1}) \\
& = P(Spam) \times P(W_0 \mid Spam) \times P(W_1 \mid W_0 \wedge Spam) \\
& \times \ \dots \ \times P(W_{N-1} \mid W_{N-2} \wedge \ \dots \ \wedge W_0 \wedge Spam)
\end{aligned}
$$

$$
\begin{aligned}
& P(Spam \wedge W_0 \wedge \ \dots \ \wedge W_{N-1}) \\
& = P(Spam) \times \prod_{n=0}^{N-1} P(W_n \mid Spam)
\end{aligned}
$$