

Parte 2

La Cognición como Inferencia Bayesiana

Platón



Aristóteles



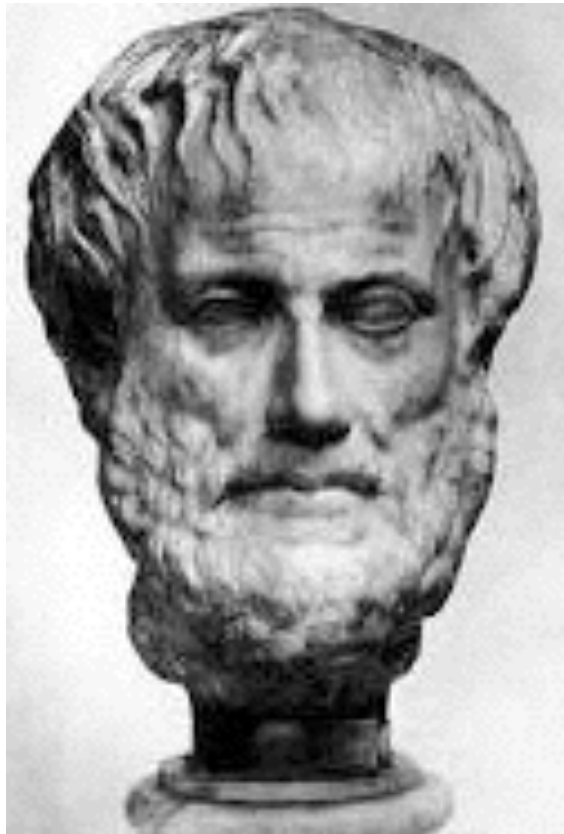
Todo **hombre** es **mortal**
Sócrates es **hombre**
Ergo **Sócrates** es **mortal**



64 silogismos pero más flexible
respecto del dominio de aplicabilidad



Kant



La Paradoja de Linda

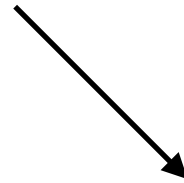
$$p(\textit{cajera}) \not\geq p(\textit{cajera} \ \& \ \textit{feminista})$$

(racional)

Sesgos y Heurísticas, hombre **irracional**
(Tversky & Kahneman, 1980s)

Juego del Dictador

Alice tiene \$100,
decide cuánto le da a Bob



Bob elige quedarse con lo que le dieron,
o nada para nadie

Racional: aceptar cualquier trato

Real: dame mínimo un ~40%..

¡Bastante independiente del monto!

Eligiendo Loterías

A.

10 premios de \$1,000,000
90 premios de \$0

B.

1 premio de \$5,000,000
8 premios de \$1,000,000
91 premios de \$0

C.

100 premios de \$1,000,000

D.

1 premio de \$5,000,000
98 premios de \$1,000,000
1 premios de \$0

$B > A$ y $C > D$... ¿racional?

\mathbb{P}

Niveles de Análisis de Marr

Inferencia Bayesiana

- Nivel *Computacional*
- Nivel *Algorítmico*
- Nivel de *Implementación*

Aproximaciones a,
Monte Carlo, etc..

Principio de Racionalidad (Anderson)

El sistema cognitivo opera todo el tiempo
para optimizar la *adaptación* del
comportamiento del organismo

Aplicando el Principio de Racionalidad

- Especificar precisamente las metas del sistema cognitivo
- Desarrollar un modelo formal del entorno al que el sistema está adaptado
- Hacer hipótesis mínimas sobre limitaciones computacionales
- Derivar el comportamiento óptimo
- Contrastar las predicciones del modelo con los experimentos
- Si las predicciones fallan, iterar

Ventajas del Principio de Racionalidad

- Construye teorías que dependen de la estructura del mundo (observable) en lugar de las mentales (no observables), a diferencia de abordajes mecanísticos
- Ofrece explicaciones de por qué los mecanismos funcionan como lo hacen, en lugar de pensar la mente como un conjunto aleatorio de postulados
- Ofrece lineamientos concretos para la construcción de teorías

¿Y qué pasó con la irracionalidad?

- Racionalidad *normativa* distinta a la *adaptativa*: *Gambler's Fallacy*
- Racionalidad *acotada* por recursos de cómputo, memoria, tiempo...
- Situación modelada distinta al entorno de adaptación

Programa de la Cognición Bayesiana

- Lógica es adecuada en contextos de **certidumbre**
- Cuando hay **incertidumbre**, el lenguaje racional es la teoría de probabilidad
- **Racionalidad acotada** por nuestros recursos de cómputo: *ilusión de irracionalidad*
- Programa **general y cuantitativo**

Tarea: armar modelos para diversos dominios...

Revisitando la Lógica

$$\text{(MP)} \quad \frac{p \Rightarrow q, p}{\therefore q}$$

$$\text{(MT)} \quad \frac{p \Rightarrow q, \neg q}{\therefore \neg p}$$

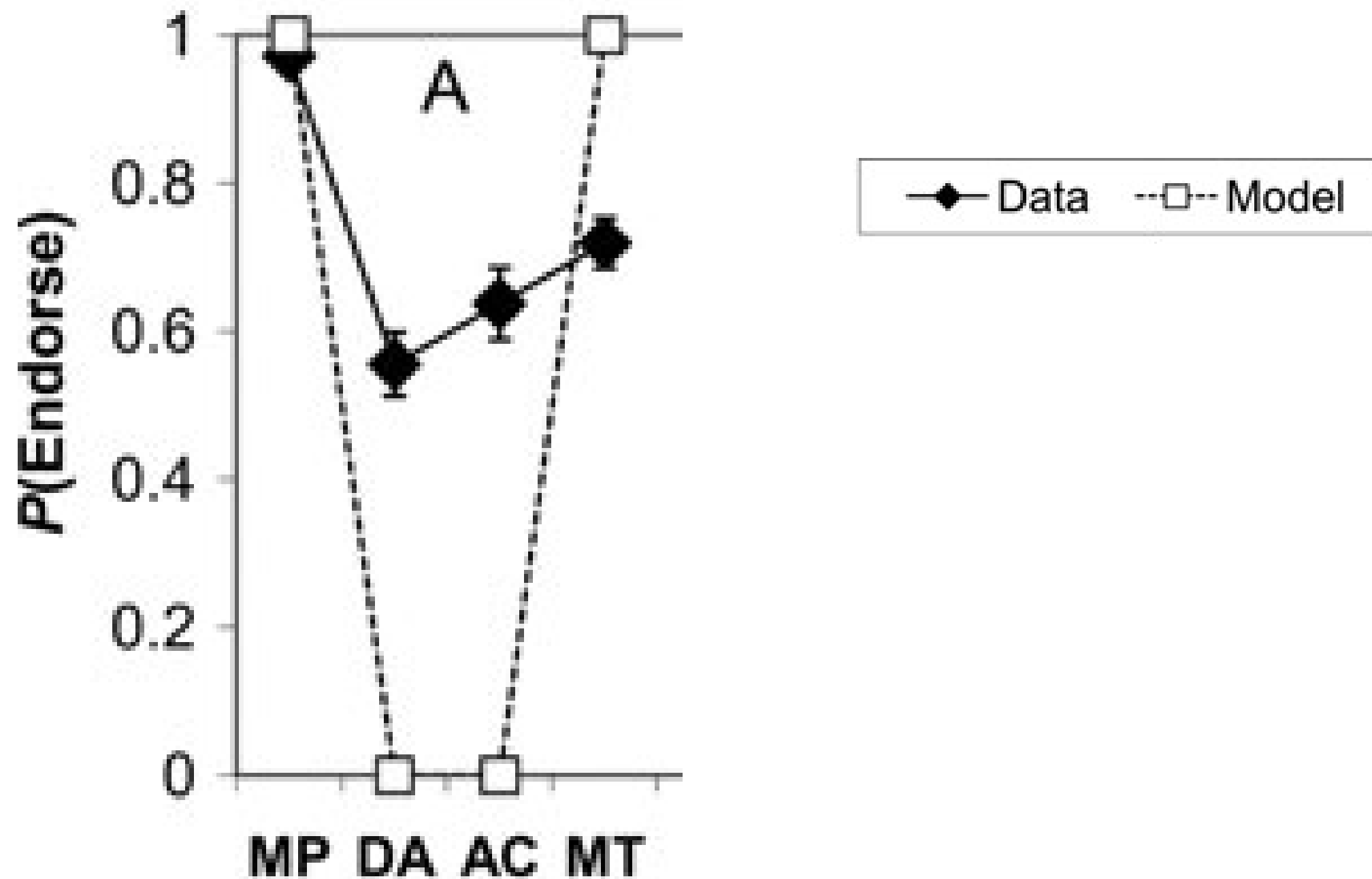


$$\text{(DA)} \quad \frac{p \Rightarrow q, \neg p}{\therefore \neg q}$$

$$\text{(AC)} \quad \frac{p \Rightarrow q, q}{\therefore p}$$



¿Qué dice la gente?



Modelo Bicondicional

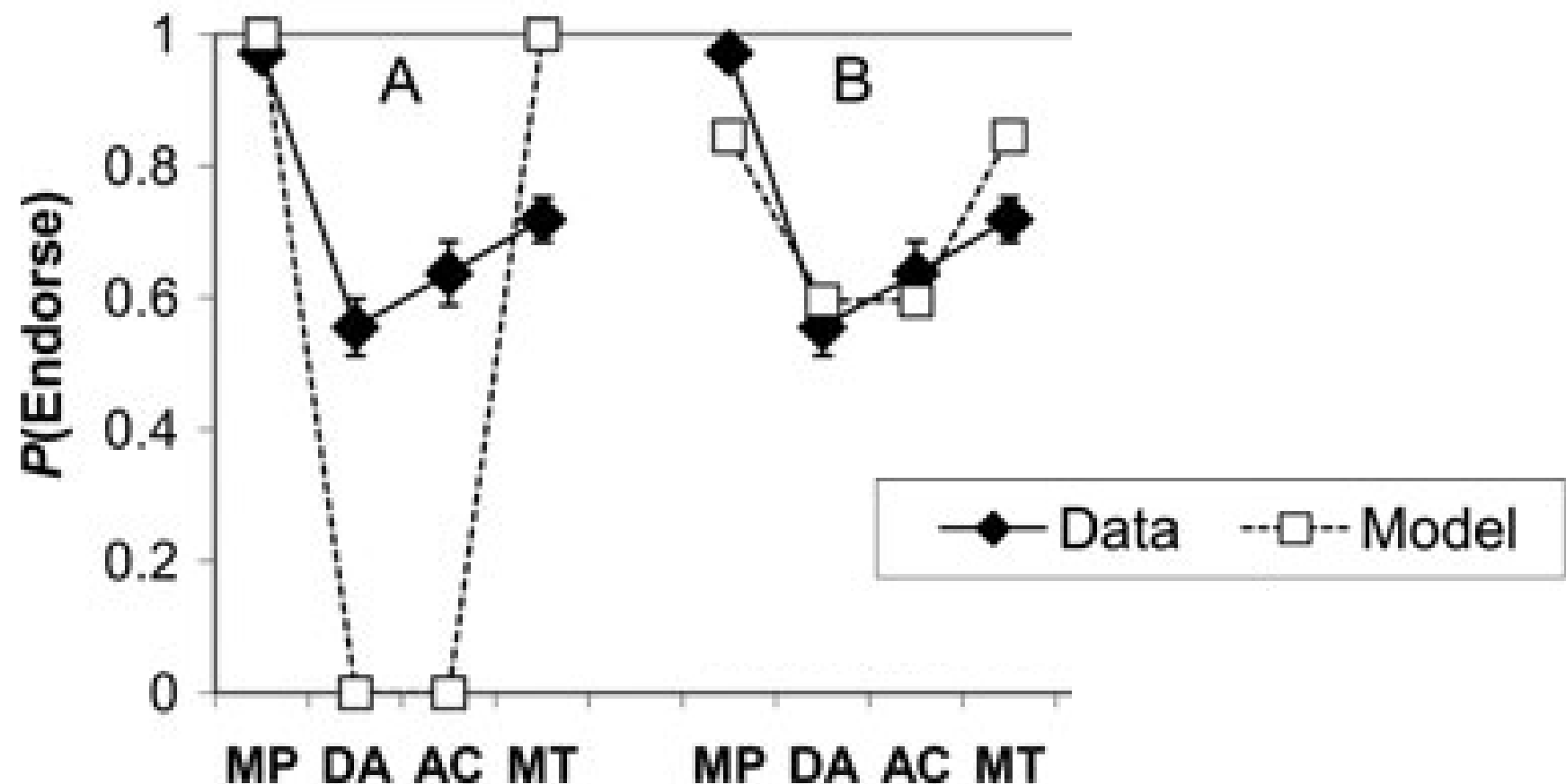
Interpretando \Rightarrow como \Leftrightarrow

Si un ave es un cisne, entonces es blanca

Si un ave es blanca, entonces es un cisne

Si me regás las plantas, entonces te pago \$20

¡Pero no te pago si no las regás!



Enfoque Bayesiano

$$P(\text{si } p \text{ entonces } q) = P(q|p)$$

¡La probabilidad de un condicional es la probabilidad condicional!

Modus Ponens

$$\text{(MP)} \quad \frac{p \Rightarrow q, p}{\therefore q} \quad \begin{array}{l} p = \text{está soleado} \\ q = \text{Juan juega al tenis} \end{array}$$

“si está soleado, Juan juega al tenis ”

$$P(q|p) = 0.9$$

“está soleado”

$$P(p)=1$$

$$\boxed{P(q) = P(q|p) = 0.9}$$

Si interpretamos el condicional como cierto (lógica): $P(q) = 1$

Modus Tollens

$$\text{(MT)} \quad \frac{p \Rightarrow q, \neg q}{\therefore \neg p} \quad \begin{array}{l} p = \text{no hay un terremoto} \\ q = \text{voy a trabajar} \end{array}$$

“no fui a trabajar” \Rightarrow “¡hay un terremoto! (no no hay)”

$$P(\neg p) = P(\neg p | \neg q) = ? \quad \text{sólo tenemos } P(q|p)$$

$$\text{Bayes: } P(\neg p | \neg q) = \dots = 1 - (1 - P(q|p)) P_0(p) / P_0(\neg q) \quad \mathbb{P}$$

$$\text{Números: } P_0(p) = 0.999; P_0(q) = 0.95; P(q|p) = 0.96\dots$$

$$\boxed{P(\neg p) = 0.2}$$

Las probabilidades están ligadas, ajustándolas:

