TOPICS IN COGNITIVE SCIENCE

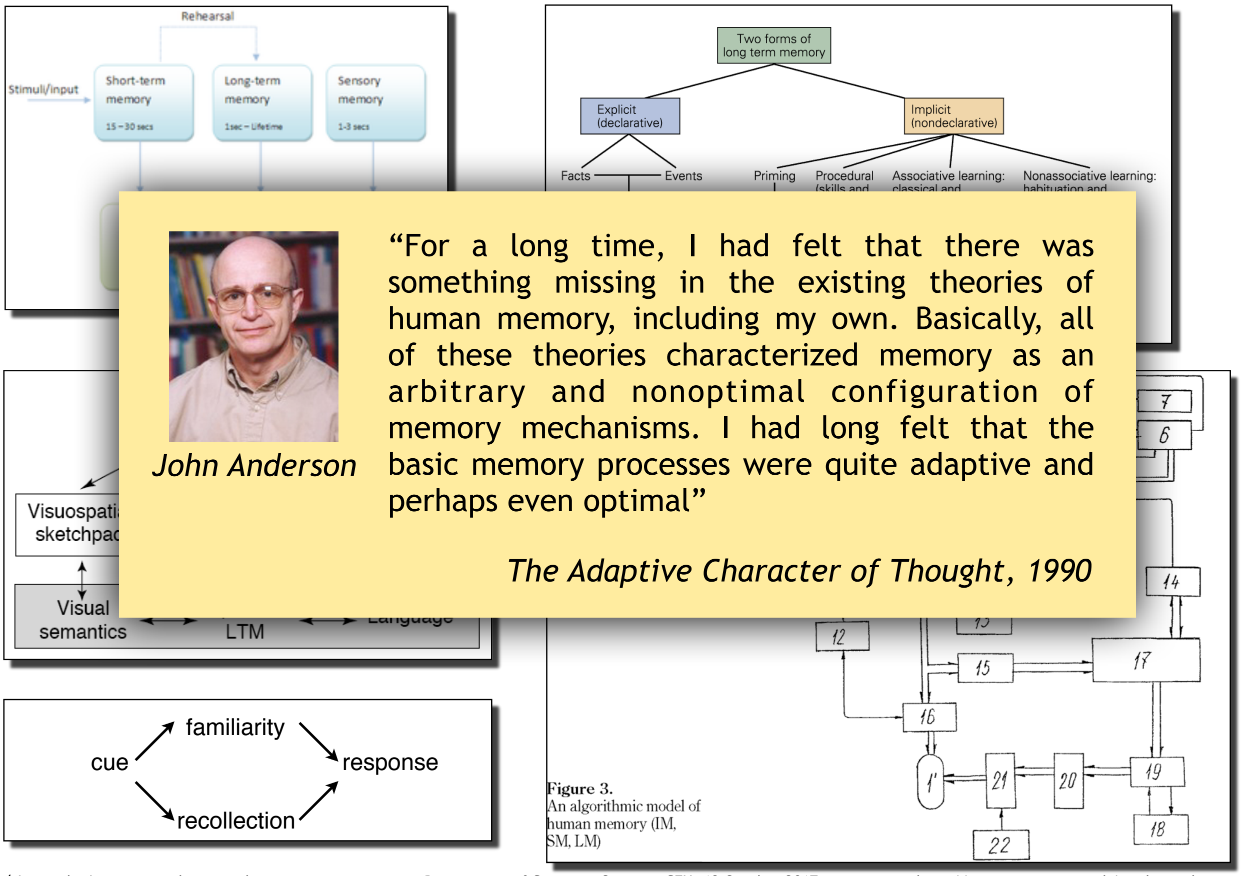
# *LEARNING: A THEORETICAL PERSPECTIVE*

MÁTÉ LENGYEL

Computational and Biological Learning Lab
Department of Engineering
University of Cambridge

Department of Cognitive Science
Central European University

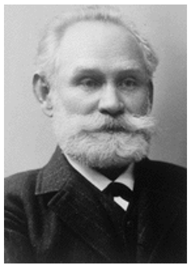
## MULTIPLE MEMORY SYSTEMS




"For a long time, I had felt that there was something missing in the existing theories of human memory, including my own. Basically, all of these theories characterized memory as an arbitrary and nonoptimal configuration of memory mechanisms. I had long felt that the basic memory processes were quite adaptive and perhaps even optimal"

*The Adaptive Character of Thought, 1990*

# PAVLOVIAN CONDITIONING



Ivan Pavlov
Nobel Prize 1904

| **before training** | **training** | **after training** |
| --- | --- | --- |
| CS → no response<br>US → response | CS+US | CS → response |

CS: bell
US: food
response: salivation
↓
**prediction!**



# THE RESCORLA-WAGNER RULE

*Rescorla & Wagner, 1972*

predict **rewards** based on stimuli

response: prediction of US

$$r = \sum_i s_i \, w_i \longrightarrow u$$

$s_i$: CS$_i$ — 0: absent, 1: present

$w_i$: ‘weight’$_i$ — degree of association of CS$_i$ with US

$u$: US — 0: absent, 1: present

error: $$E = (u - r)^2 = \left(u - \sum_i s_i \, w_i\right)^2$$

minimise error wrt. weights
‘stochastic gradient descent’

$$-\frac{\partial E}{\partial w_i} \propto \underbrace{(u - r)}_{\delta} \, s_i$$

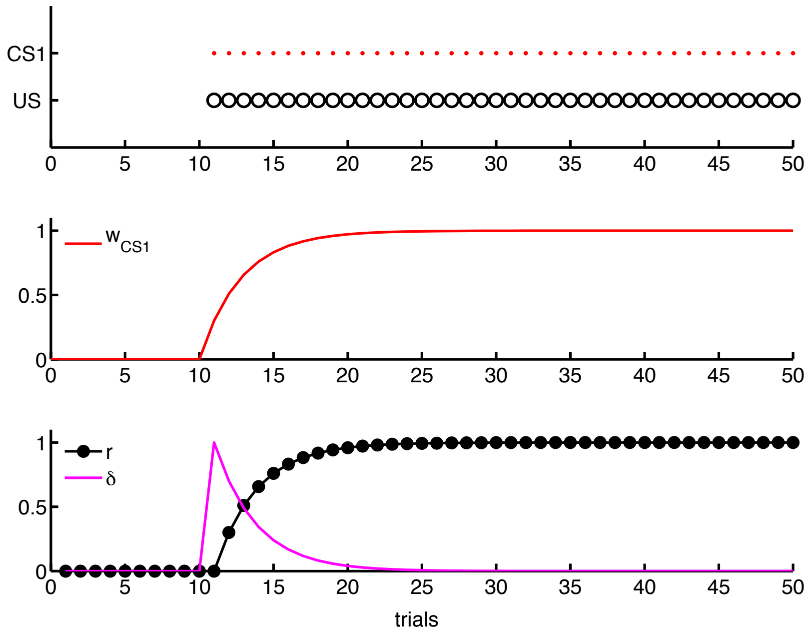
signed prediction error

on each trial update: $$w_i \to w_i + \epsilon \, (u - r) \, s_i$$

$\epsilon$: learning speed ≪ 1

# PAVLOVIAN CONDITIONING REVISITED

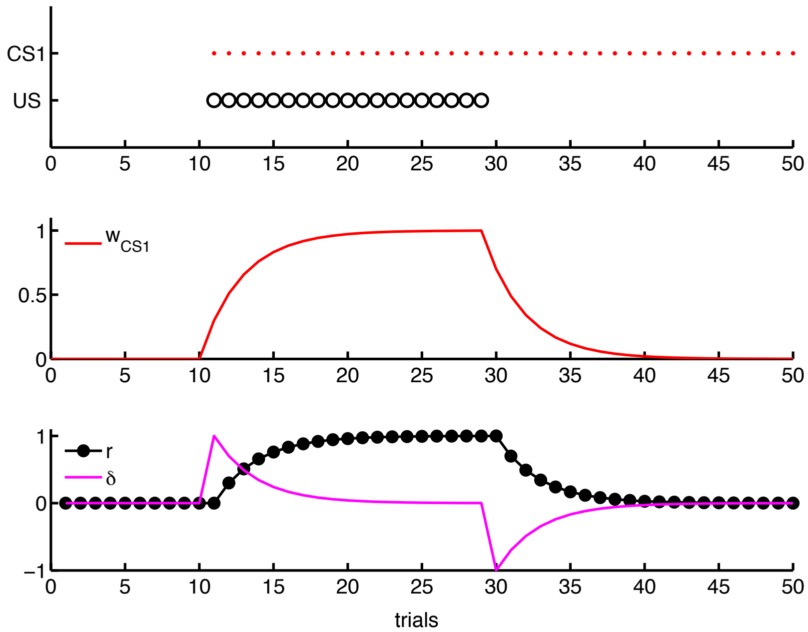
| before training | training | after training |
|---|---|---|
| CS →no response<br>US →response | CS+US | CS → response |





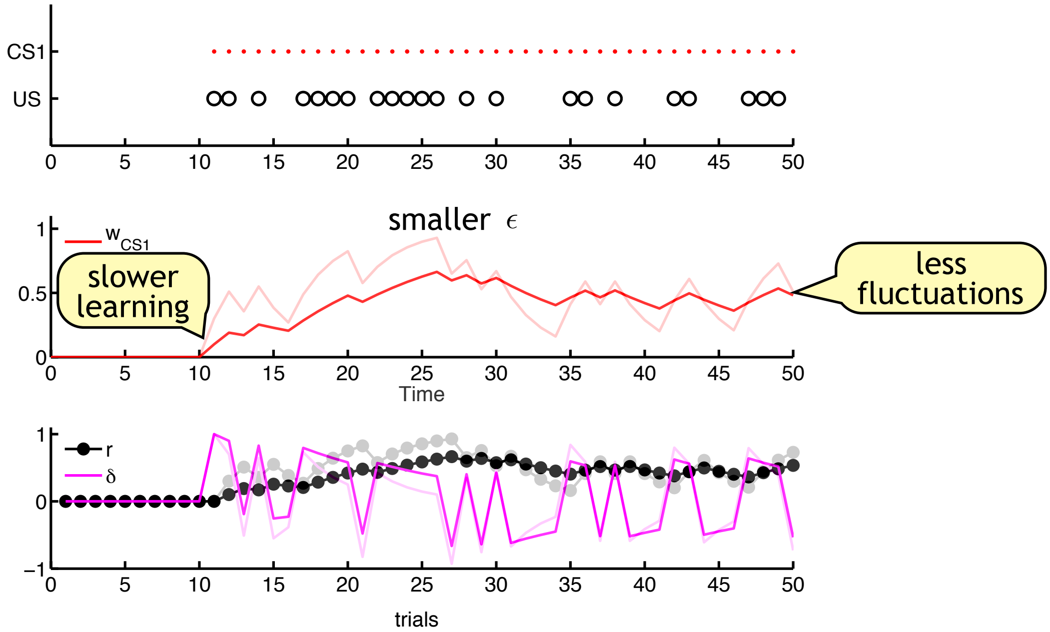

# PAVLOVIAN EXTINCTION

| before training | training | | after training |
|---|---|---|---|
| | phase 1 | phase 2 | |
| CS →no response<br>US →response | CS+US | CS | CS → no response |



# PARTIAL REINFORCEMENT

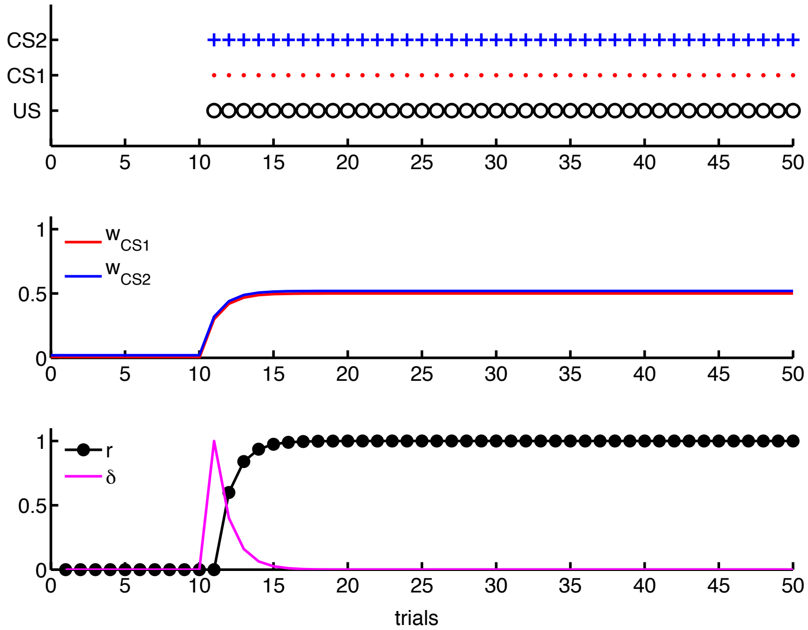
| before training | training | after training |
|---|---|---|
| CS →no response<br>US →response | CS, CS+US | CS → weak response |





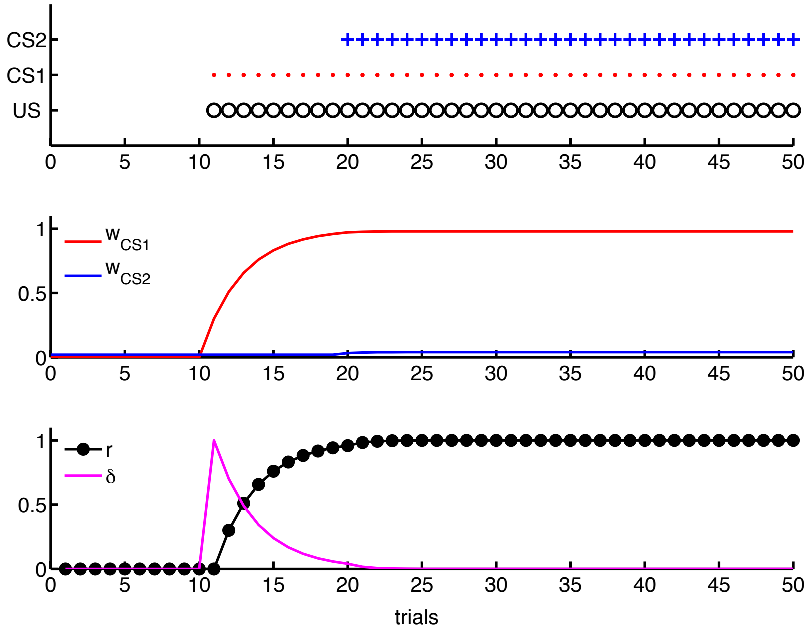

# OVERSHADOWING

| before training | training | after training |
|---|---|---|
| $CS_1$ →no response<br>$CS_2$ →no response<br>US →response | $CS_1$+$CS_2$+US | $CS_1$ →weak response<br>$CS_2$ →weak response |



# BLOCKING

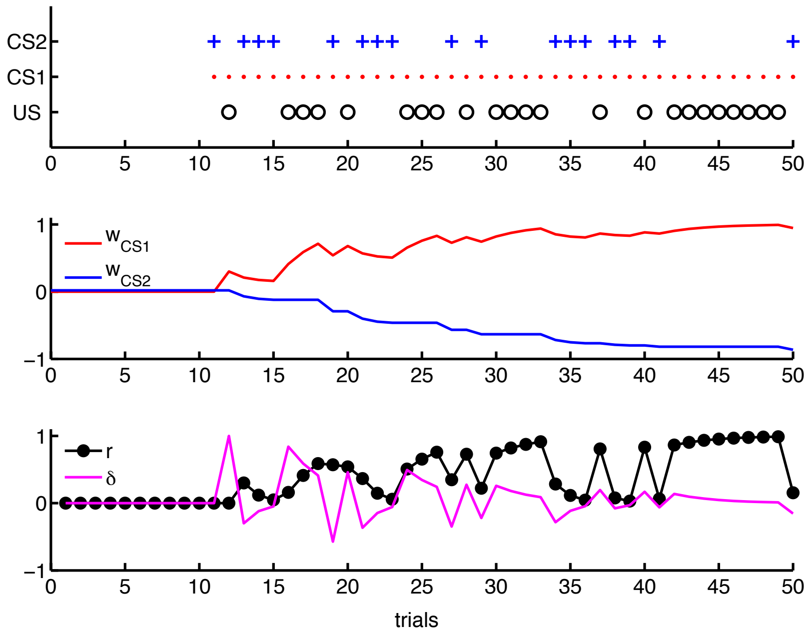
| before training | training | | after training |
|---|---|---|---|
| | phase 1 | phase 2 | |
| $CS_1$ →no response<br>$CS_2$ →no response<br>US →response | $CS_1$+US | $CS_1$+$CS_2$+US | $CS_1$ →response<br>$CS_2$ →no response |





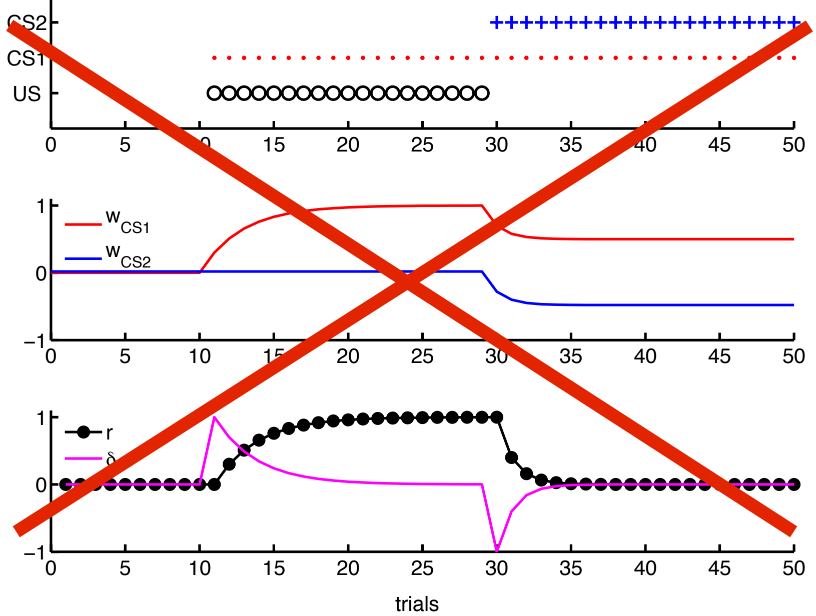

# INHIBITORY CONDITIONING

| before training | training | after training |
|---|---|---|
| $CS_1$ →no response<br>$CS_2$ →no response<br>US →response | $CS_1$+US<br>$CS_1$+$CS_2$ | $CS_1$ →response<br>$CS_2$ →no response (inhibition) |



# SECONDARY CONDITIONING

| before training | training phase 1 | training phase 2 | after training |
|---|---|---|---|
| $CS_1$ →no response<br>$CS_2$ →no response<br>US →response | $CS_1$+US | $CS_1$+$CS_2$ | $CS_2$ →response |



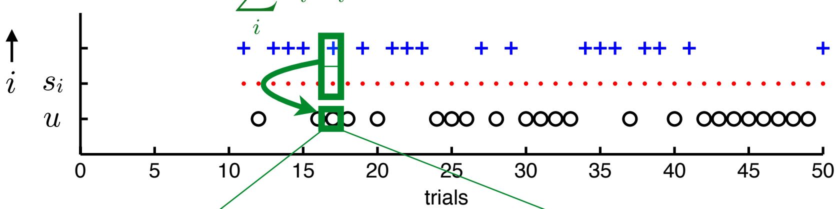



# FROM RESCORLA-WAGNER TO TEMPORAL-DIFFERENCE LEARNING

**RW** predict **immediate rewards** based on **current stimuli**

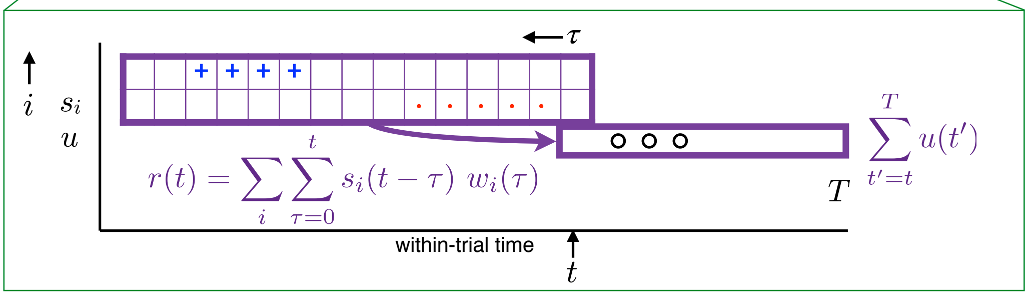
$$w_i \to w_i + \epsilon\,(u - r)\,s_i$$

$$r = \sum_i s_i\,w_i$$

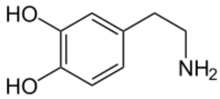



**TD** consider within-trial time
predict **total future rewards** based on **stimulus history**

$$w_i(\tau) \to w_i(\tau) + \epsilon\,[u(t) + r(t+1) - r(t)]\,s_i(t-\tau)$$



# NEURAL SUBSTRATE: DOPAMINE



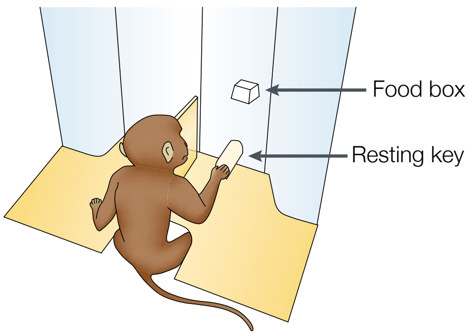

only metabotropic receptors
(eg. acting on adenylyl cyclase)

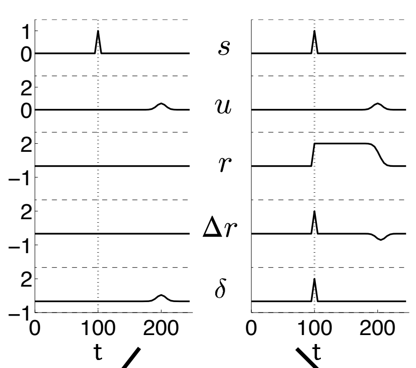
- drugs: cocaine, amphetamine → high dopamine levels
- disorders: schizophrenia, Parkinson's disease, ADHD
- implicated in self-stimulation, addiction

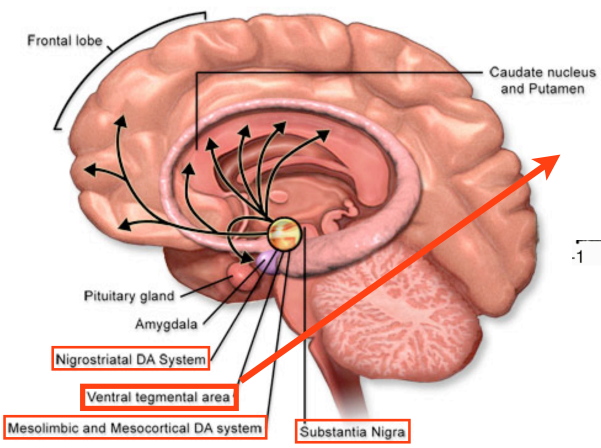


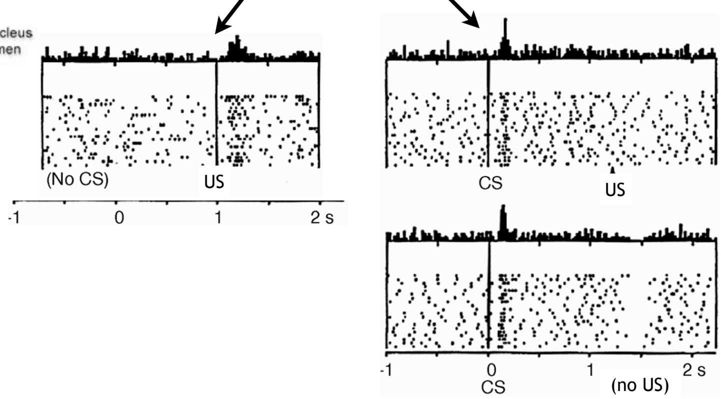
# DOPAMINE = PREDICTION ERROR

*Schultz, Dayan, and Montague, Science 1997*



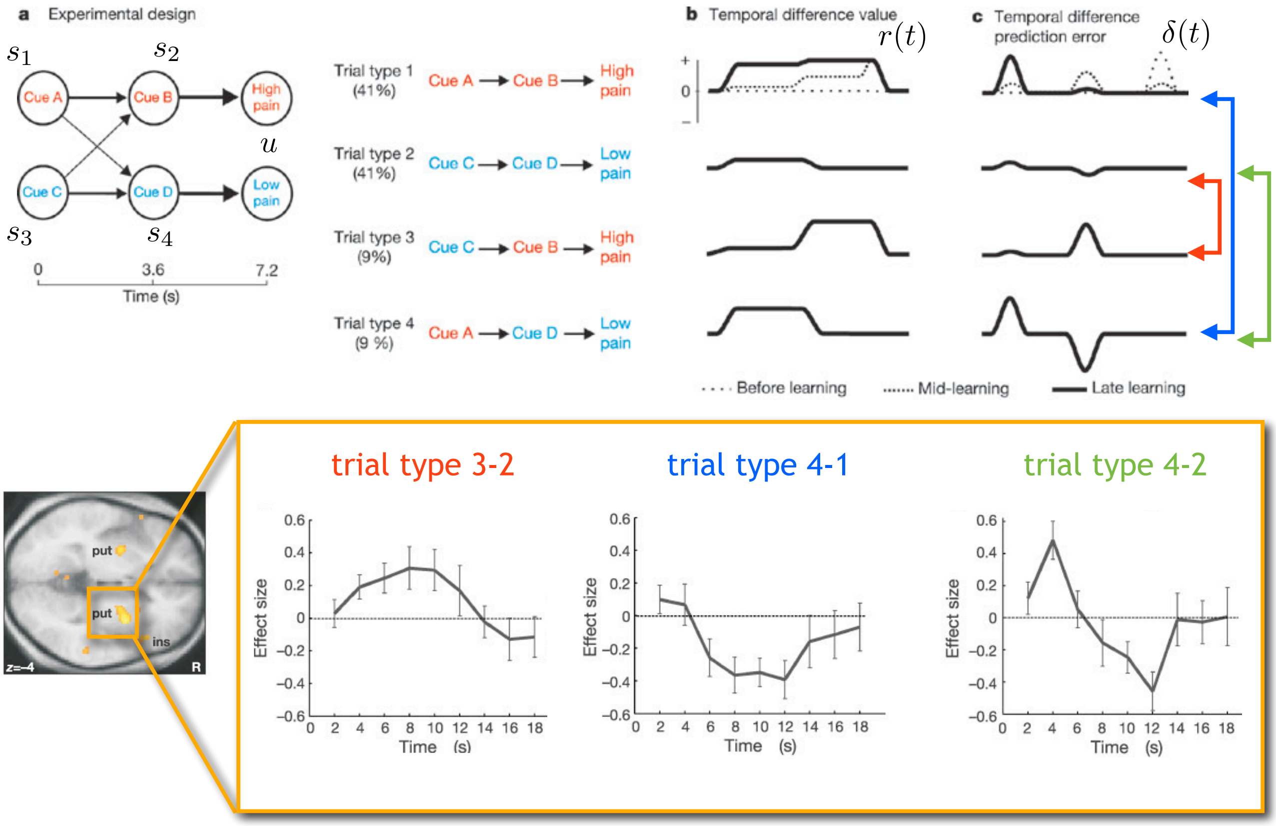



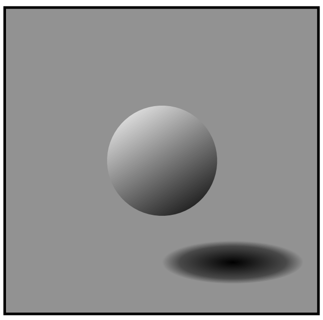
# HIGHER-ORDER LEARNING IN HUMANS

*Seymour et al, Nature 2004*






# PROBABILISTIC MODELS



**“perception is unconscious inference”**
**& memory & learning & …**

*hidden variables*

$y_1$ position of lighting source

$y_2$ convexity of shape

shading $x_1$

shadow $x_2$

*observed variables*

*Hermann von Helmholtz*
1821-1894

There are things known and there are things unknown, and between are **the rules of probability**

product: $\mathrm{P}(X,Y) = \mathrm{P}(Y,X) = \mathrm{P}(X|Y)\ \mathrm{P}(Y)$

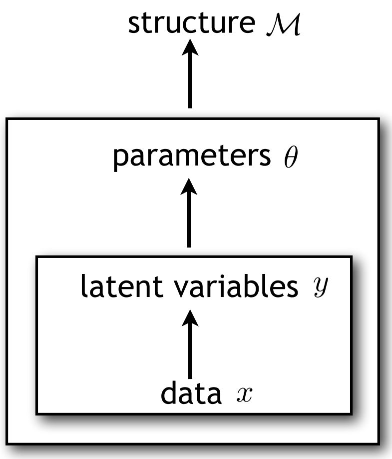
⇒ Bayes’ rule: $\mathrm{P}(Y|X) = \dfrac{\mathrm{P}(X|Y)\ \mathrm{P}(Y)}{\mathrm{P}(X)}$

$\mathrm{P}(X,Y) = \mathrm{P}(X)\ \mathrm{P}(Y)$ iff $X$ and $Y$ are independent!

sum: (marginalisation) $\mathrm{P}(X) = \sum_Y \mathrm{P}(X,Y)$

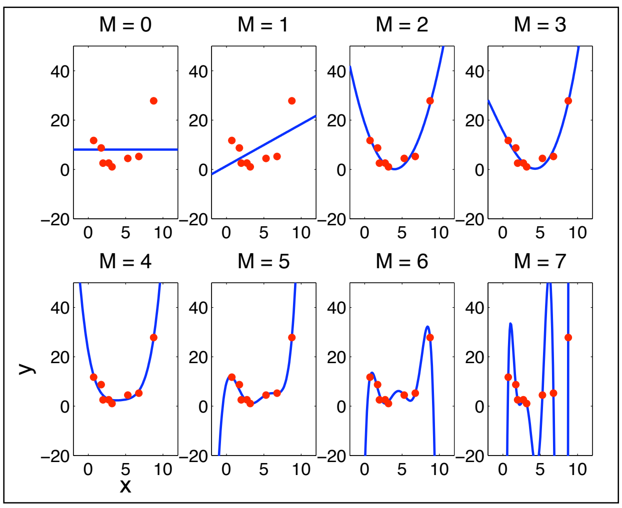
*Rev. Thomas Bayes*
1702-1761

# PROBABILISTIC INFERENCE AND LEARNING






# MODEL COMPARISON




*courtesy of Zoubin Ghahramani*

- under which model do I get the best fit?

$$\mathrm{P}\left(\mathcal{D}|\hat{\theta}_{\mathrm{ML}}, \mathcal{M}\right)$$

parameters — model structure

what is the likelihood of the model with *the best* parameters?

overfitting!

- which model has the highest likelihood?

$$\mathrm{P}(\mathcal{D}|\mathcal{M}) = \sum_{\theta} \mathrm{P}(\mathcal{D}|\theta, \mathcal{M})\, \mathrm{P}(\theta|\mathcal{M})$$

what is the average likelihood of the model with *randomly chosen* parameters?

# BAYESIAN MODEL SELECTION

Rev. Bayes

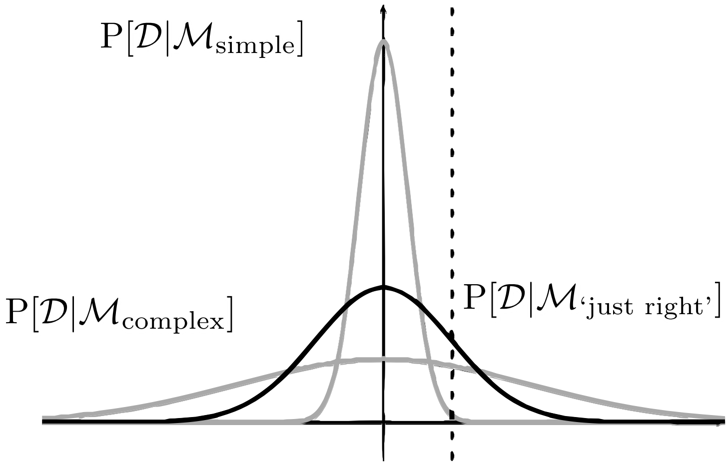
**a model defines a probability distribution over data sets**

$\mathrm{P}[\mathcal{D}|\mathcal{M}]$

data coming from the environment — model stored in memory

Occam's razor

'entities should not be multiplied beyond necessity'

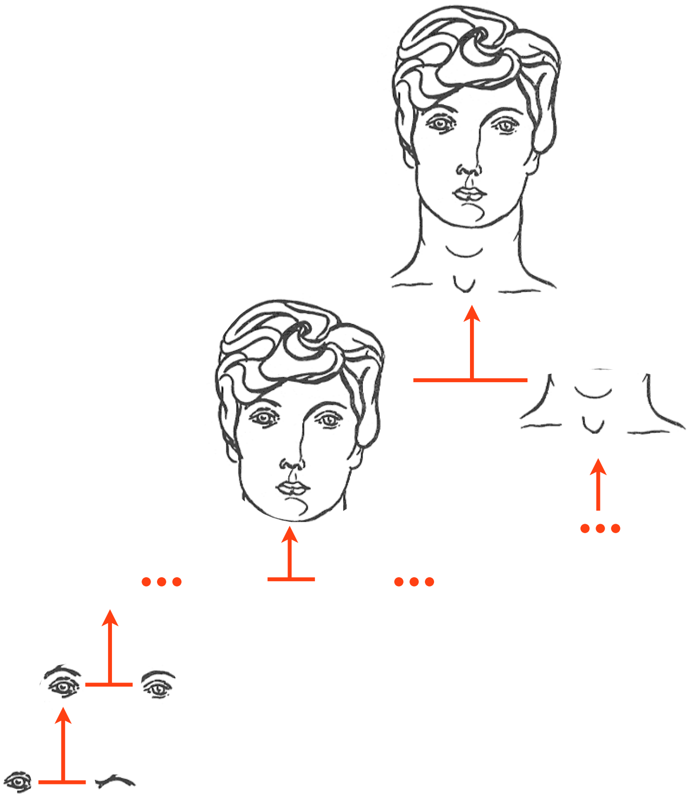



possible data sets $\mathcal{D}$

experienced data set $\mathcal{D}^*$



# CHUNK LEARNING: HIERARCHICAL PAIR-WISE ASSOCIATIVE?



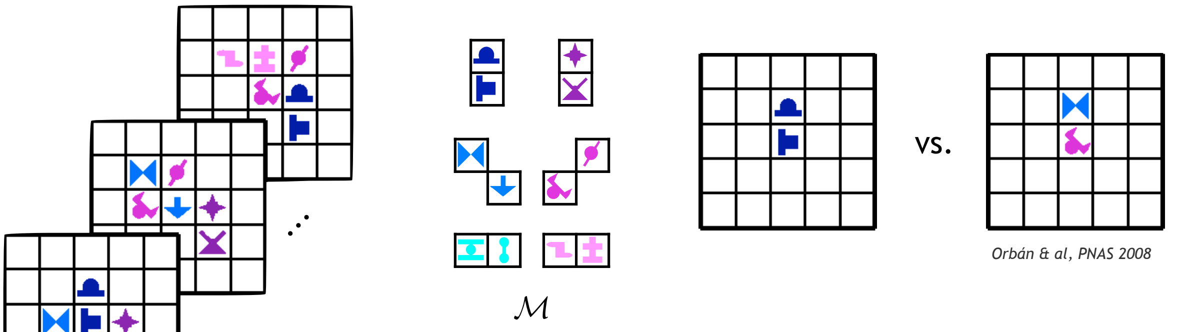
Gergő Orbán

József Fiser

Richard Aslin

# VISUAL PATTERN LEARNING

**familiarization** 'pay attention' | **inventory** | **test** 'which one looks more familiar?'



*Orbán & al, PNAS 2008*

$$\mathcal{D} = \text{scene}_1, \ldots \text{scene}_n \rightarrow \hat{\mathcal{M}} = \underset{\mathcal{M}}{\text{argmax}}\, \text{P}(\mathcal{M}|\mathcal{D}) \rightarrow \text{P}\left(\text{scene}_\text{A}|\hat{\mathcal{M}}\right) \text{ vs. } \text{P}\left(\text{scene}_\text{B}|\hat{\mathcal{M}}\right)$$

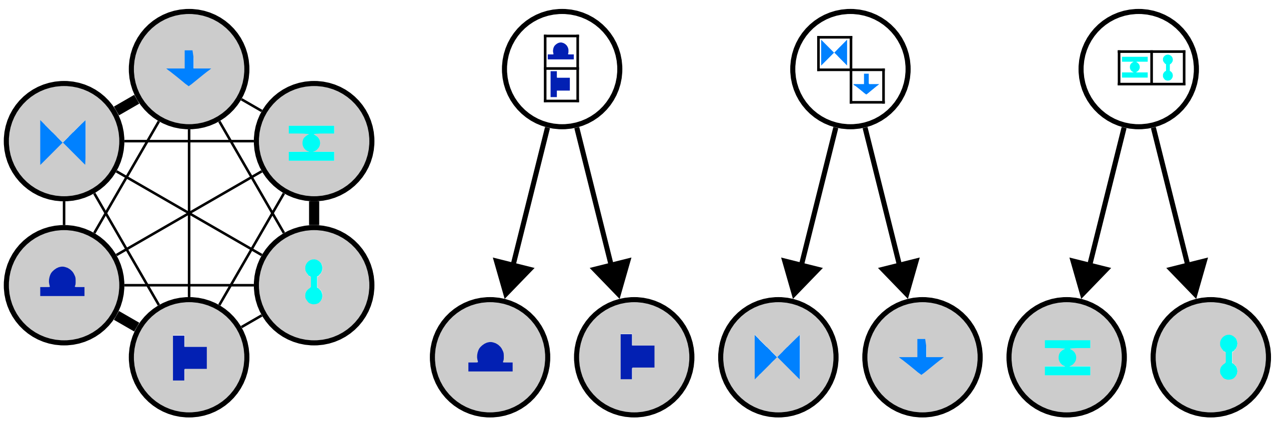
**how do humans learn a statistical model of their environment?**

- associative learning (fitting 2nd order max-entropy model)
- Bayesian model selection (inferring hidden causal structure)



# ALTERNATIVE THEORIES

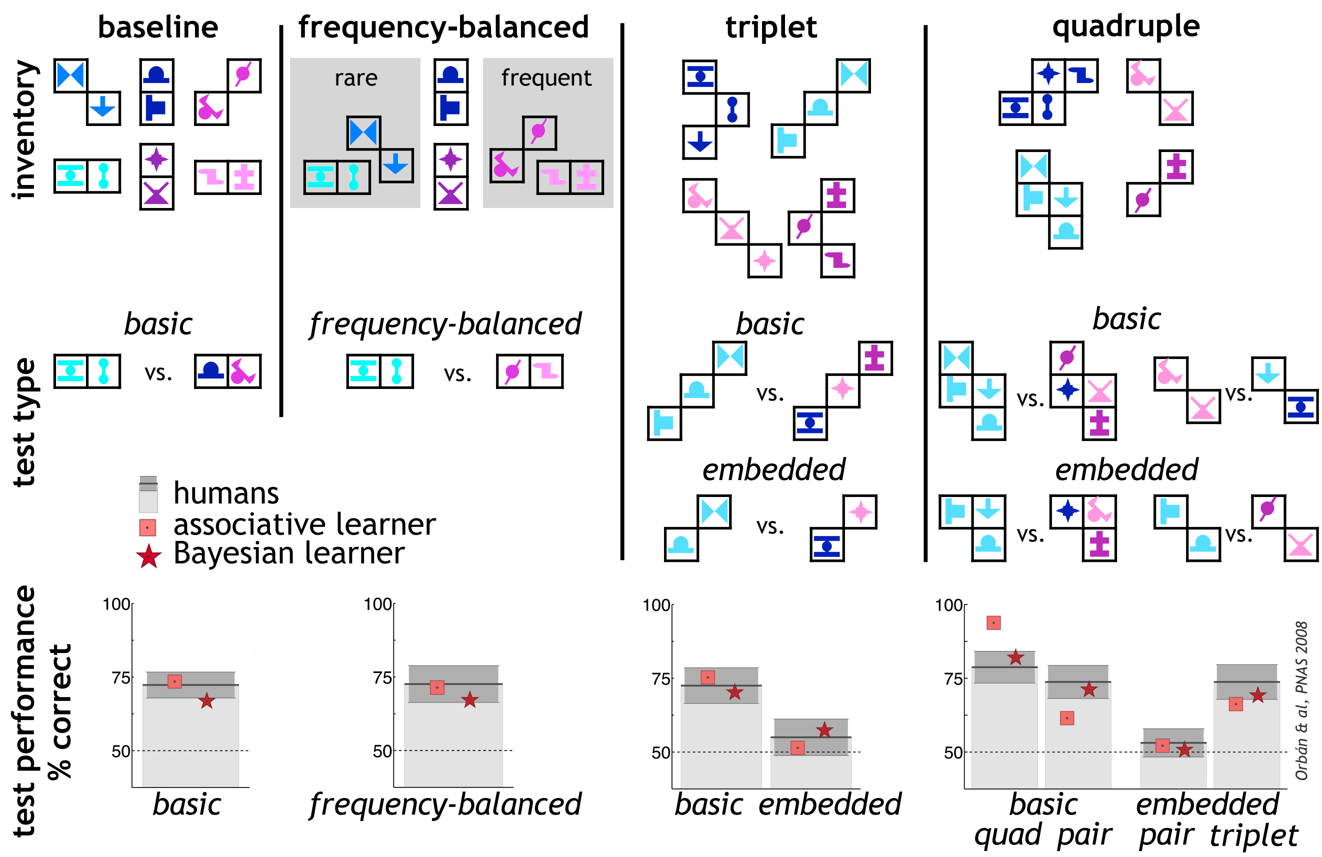
**associative learning** | **Bayesian learning**


*Orbán & al, PNAS 2008*

Boltzmann machine
+ Gaussian Markov random field

sigmoid belief network
+ product of (conditional) Gaussian experts

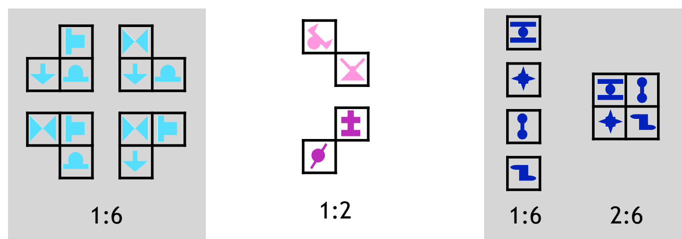
# MULTIPLE EXPERIMENTS





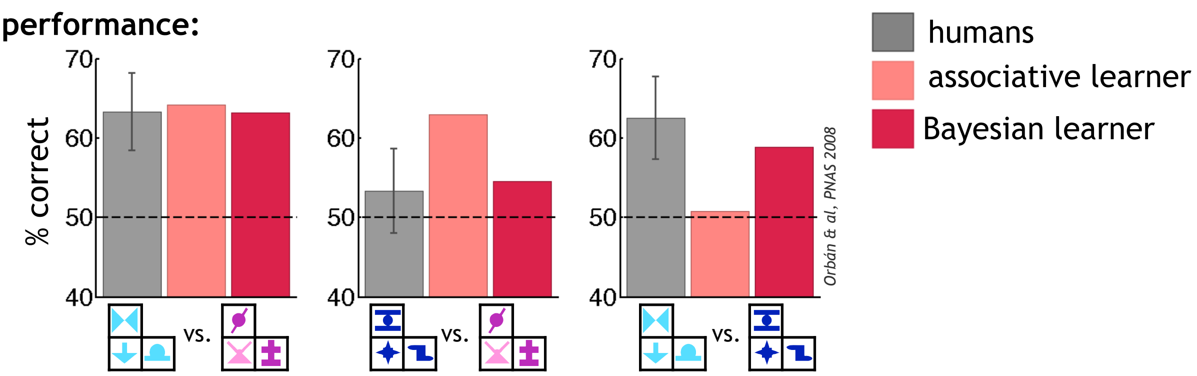

# ASSOCIATIVE VS. BAYESIAN LEARNING

**inventory:**



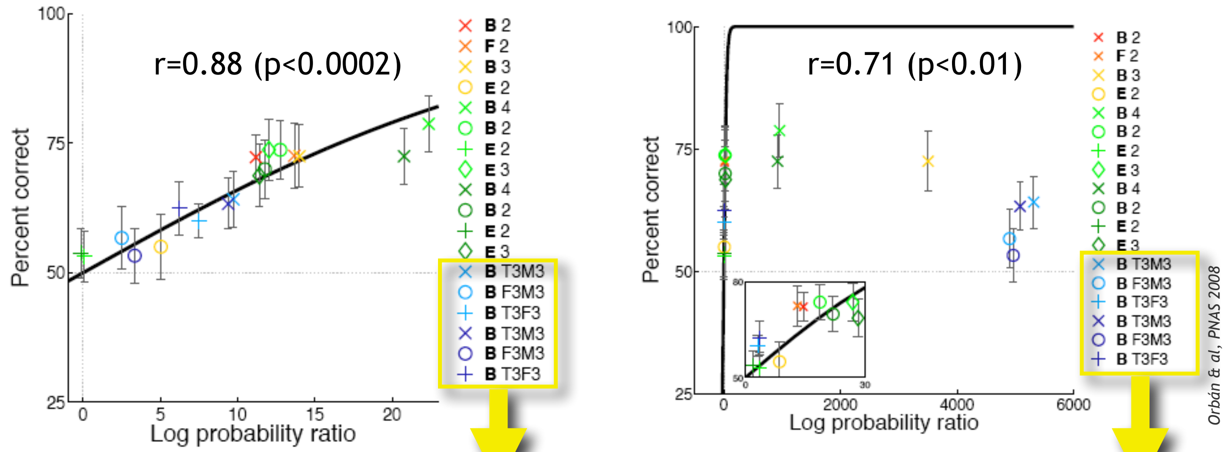

| | | | |
|---|---|---|---|
| 1st order statistic: shape frequencies | | 3 × 1/6 | 1/6 + 2/6 |
| 2nd order statistic: pairwise correlations | both present | 2 × 1/6 | 2/6 |
| | both absent | 2/6 | 2 × 1/6 |
| | one present | 2 × 1/6 | 2 × 1/6 |

**test performance:**



# QUANTITATIVE COMPARISON

Bayesian learner | associative learner

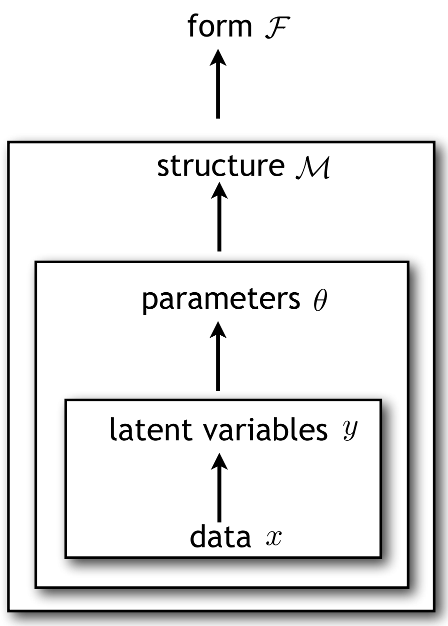



predictions without further fitting

r=0.92 (p<0.006) | r=−0.23 (p>0.65)



# PROBABILISTIC INFERENCE AND LEARNING



# GOING UP, UP, UP …

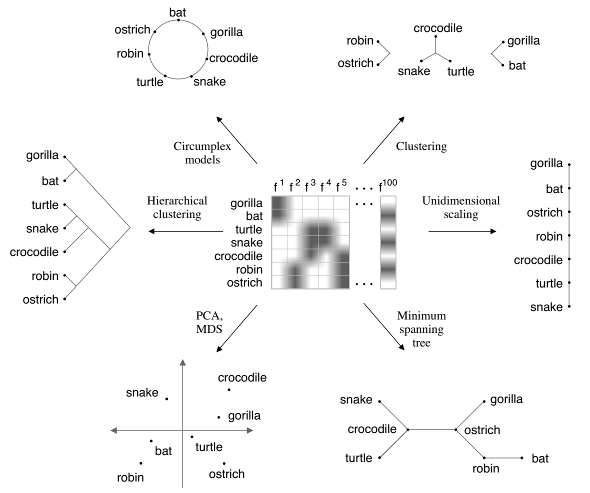
## why constrain ourselves to one model form?

form $\mathcal{F}$
↓
structure $\mathcal{M}$
↓
parameters $\theta$
↓
latent variables $y$
↓
data $x$

$$\mathrm{P}(x|\mathcal{F}) = \sum_{\mathcal{M}} \mathrm{P}(x|\mathcal{M},\mathcal{F})\,\mathrm{P}(\mathcal{M}|\mathcal{F}) \longrightarrow \hat{\mathcal{F}} = \operatorname*{argmax}_{\mathcal{F}} \mathrm{P}(x|\mathcal{F})$$

$$\mathrm{P}(x|\mathcal{M}) = \sum_{\theta} \mathrm{P}(x|\theta,\mathcal{M})\,\mathrm{P}(\theta|\mathcal{M}) \longrightarrow \hat{\mathcal{M}} = \operatorname*{argmax}_{\mathcal{M}} \mathrm{P}(x|\mathcal{M})$$

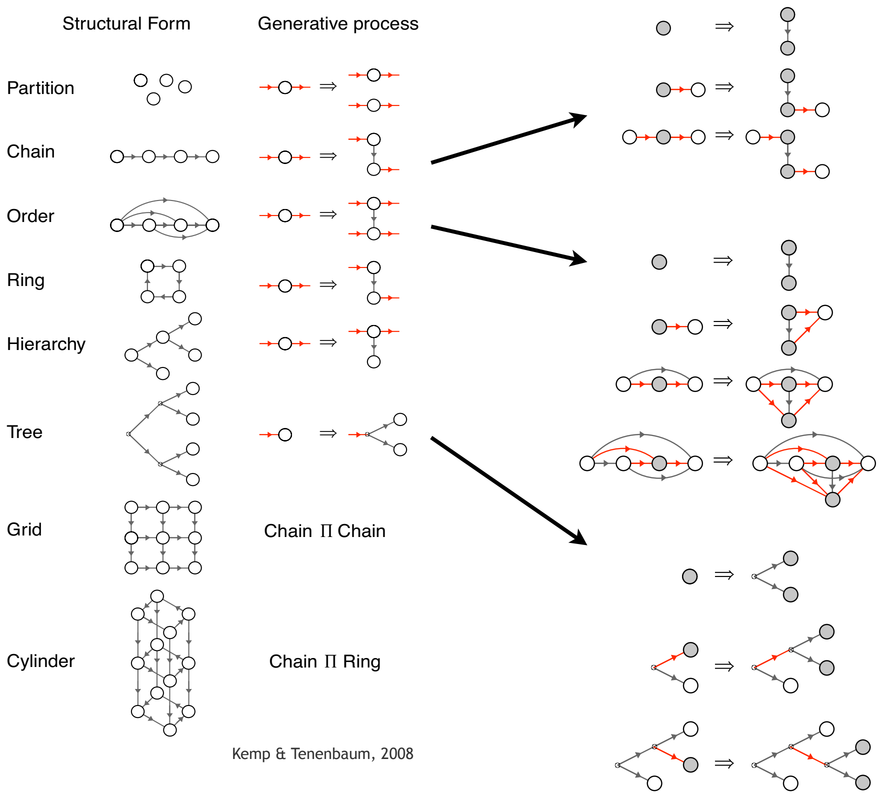
$$\mathrm{P}(x|\theta) = \sum_{y} \mathrm{P}(x|y,\theta)\,\mathrm{P}(y|\theta) \longrightarrow \hat{\theta} = \operatorname*{argmax}_{\theta} \mathrm{P}(x|\theta)$$




*Kemp & Tenenbaum, 2008*

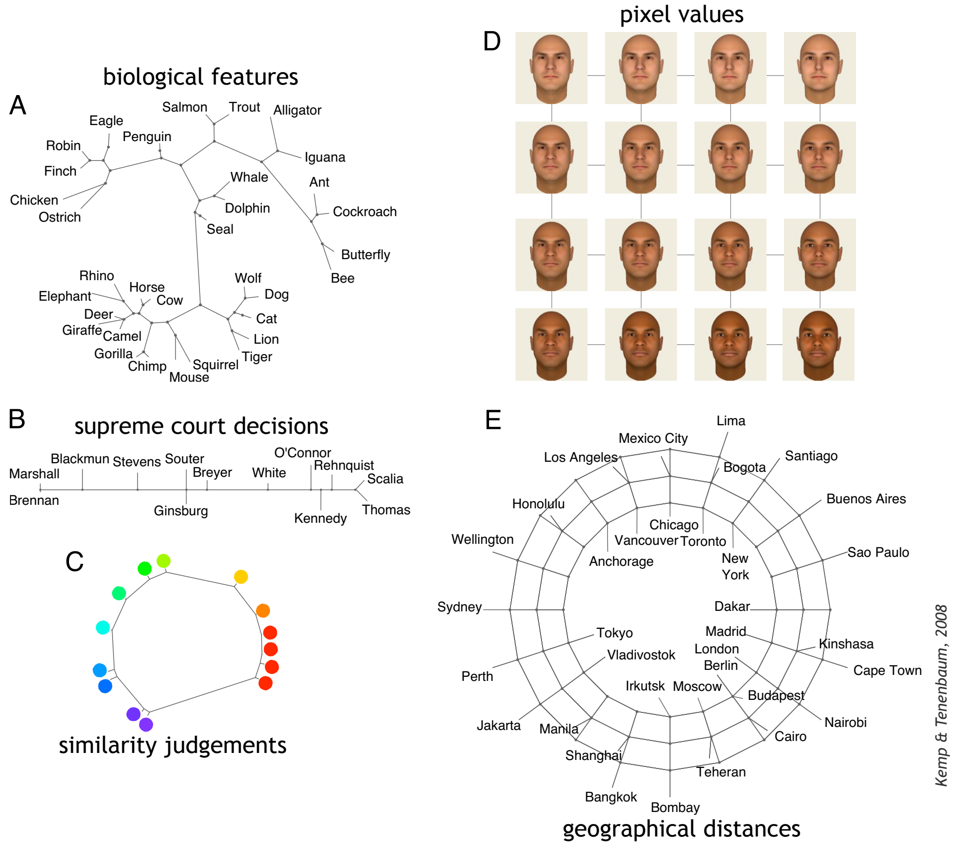


# GRAPH GRAMMARS



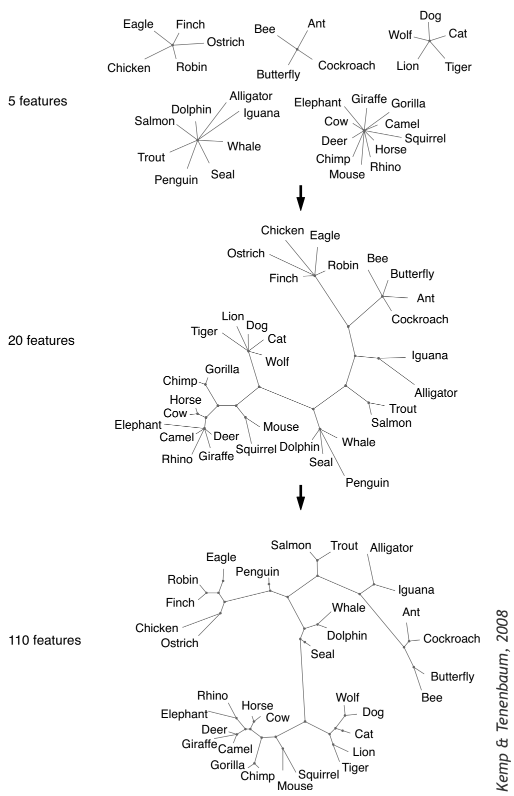

Kemp & Tenenbaum, 2008

# DISCOVERING STRUCTURAL FORM






# LEARNING STRUCTURAL FORM

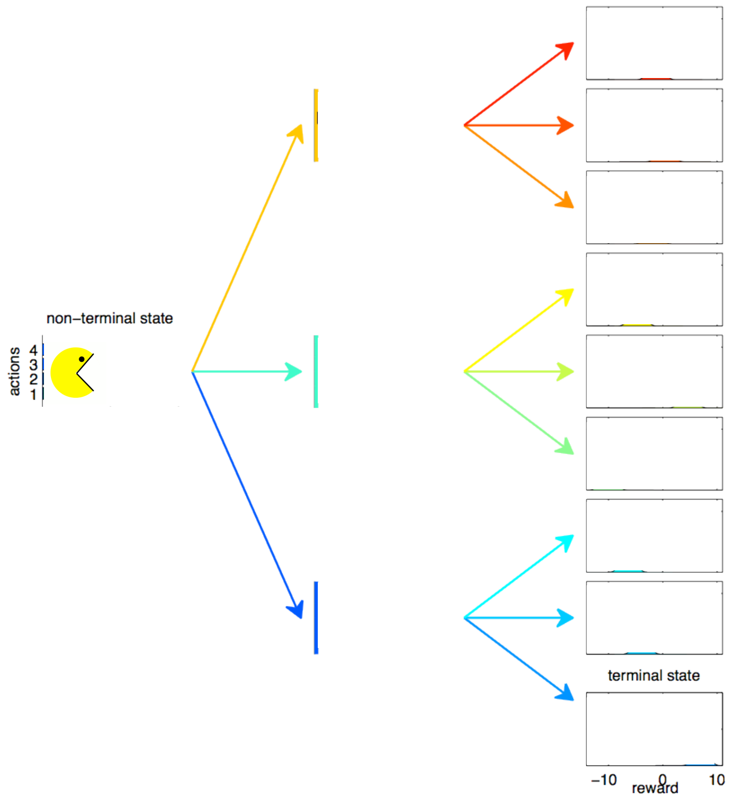


# PROBABILISTIC INFERENCE AND LEARNING

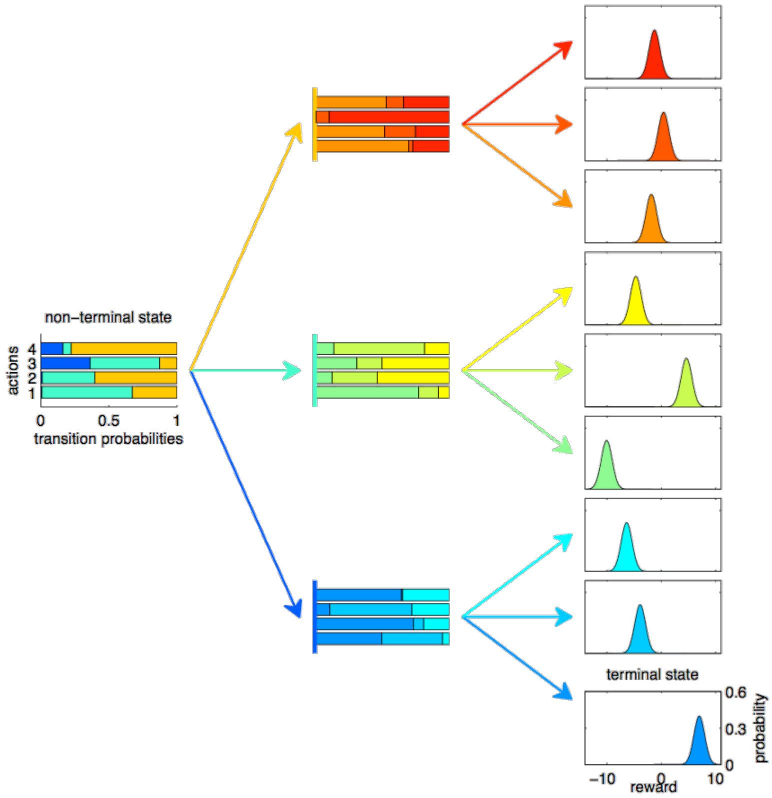
| cognitive science: *theory* | cognitive science: *experiments* | | neuroscience: *theory* | neuroscience: *experiments* |
|---|---|---|---|---|
| | | form $\mathcal{F}$ | | |
| ✓ | ?✓ | ↑ | ? | ? |
| | | structure $\mathcal{M}$ | | |
| ✓ | ✓ | ↑ | ? | ? |
| | | parameters $\theta$ | | |
| ✓ | ✓ | ↑ | ✓✓ | ?✓ |
| | | latent variables $y$ | | |
| ✓ | ✓ | ↑ | ✓✓✓ | ?✓ |
| | | data $x$ | | |



# LIFE = *SEQUENTIAL* DECISION-MAKING UNDER RISK AND *UNCERTAINTY*



# LIFE = *SEQUENTIAL* DECISION-MAKING UNDER RISK AND *UNCERTAINTY*




**PRINCIPLES**

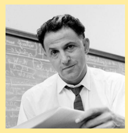
Bayes' rule
how to infer unknowns based on observations

unknowns

$$\mathrm{P}(\mathcal{Y}|\mathcal{X}) = \frac{\mathrm{P}(\mathcal{X}|\mathcal{Y})\ \mathrm{P}(\mathcal{Y})}{\sum_{\mathcal{Y}'} \mathrm{P}(\mathcal{X}|\mathcal{Y}')\ \mathrm{P}(\mathcal{Y}')}$$

observations

Bellman equation
how to select the best action



$$Q(s,a) = \sum_{s'} \mathcal{P}^{a}_{ss'} \left[\mathcal{R}_s + \gamma \max_{a'} Q(s,a)\right]$$

**+ approximations**

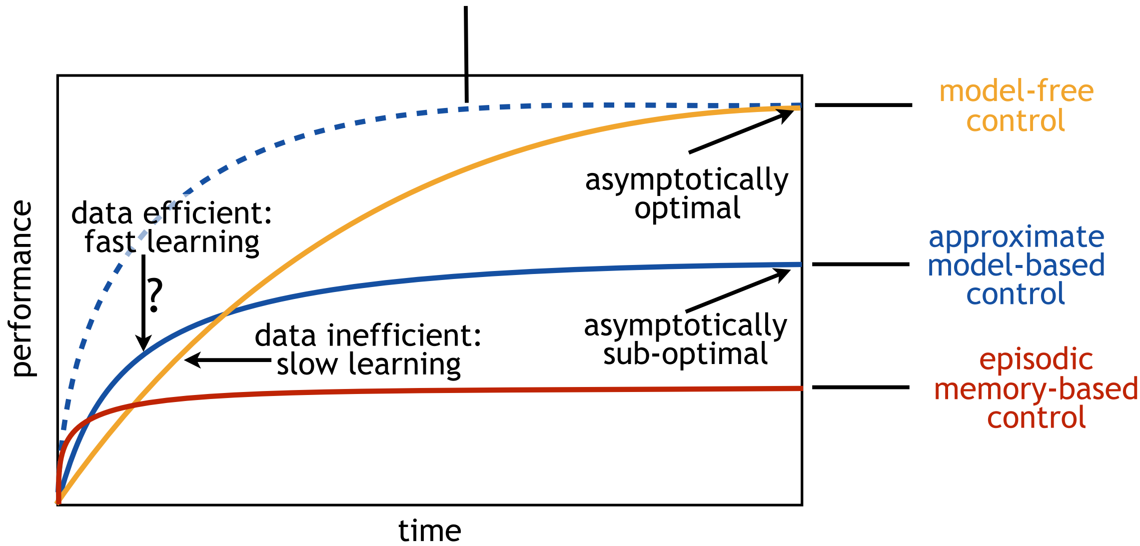


# DATA EFFICIENCY VS. COMPUTATIONAL COMPLEXITY

experience *data* —**learning**→ internal representation —**control**→ planning for the future *sequential decision making*

learning: clever / dull

control: hard / easy



# MULTIPLE MEMORY SYSTEMS: THE VIEW FROM RL