uai2011

The 27th Conference on Uncertainty in Artificial Intelligence

Active Inference and Uncertainty

Karl Friston

Abstract

In this presentation, I will rehearse the free-energy formulation of action and perception, with a special focus on the representation of uncertainty: The free-energy principle is based upon the notion that both action and perception are trying to minimise the surprise (prediction error) associated with sensory input. In this scheme, perception is the process of optimising sensory predictions by adjusting internal brain states and connections; while action is regarded as an adaptive sampling of sensory input to ensure it conforms to perceptual predictions (this is known as active inference). Both action and perception rest on an optimum representation of uncertainty, which corresponds to the precision of prediction error. Neurobiologically, this may be encoded by the postsynaptic gain of prediction error units. I hope to illustrate the plausibility of this framework using simple simulations of cued, sequential, movements. Crucially, the predictions driving movements are based upon a hierarchical generative model that infers the context in which movements are made. This means that we can temporarily confuse agents by changing the context (order) in which cues are presented. These simulations provide a (Bayes-optimal) simulation of contextual uncertainty and set-switching that can be characterised in terms of behaviour and electrophysiological responses. Interestingly, one can lesion the encoding of precision (postsynaptic gain) to produce pathological behaviours that are reminiscent of those seen in Parkinson's disease. I will use this as a toy example of how information theoretic approaches to uncertainty may help understand action selection and set-switching

UNCERTAINTY, ansen seale (2005)

“Objects are always imagined as being present in the field of vision as would have to be there in order to produce the same impression on the nervous mechanism” - Hermann Ludwig Ferdinand von Helmholtz

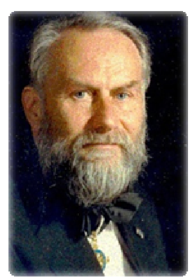
Richard Gregory

Geoffrey Hinton

Thomas Bayes

From the Helmholtz machine to the Bayesian brain and self-organization

Richard Feynman



Hermann Haken

Overview

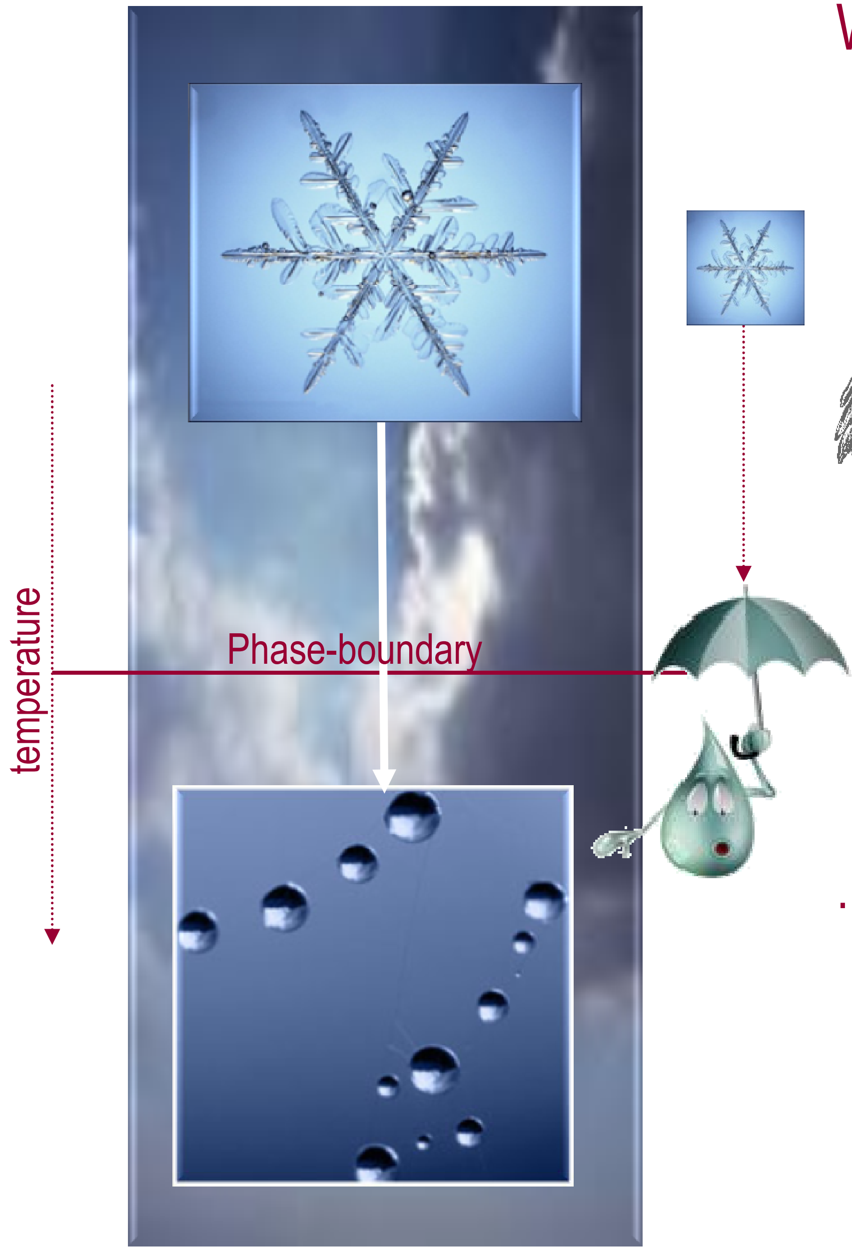
Ensemble dynamics
- Entropy and equilibria
- Free-energy and surprise

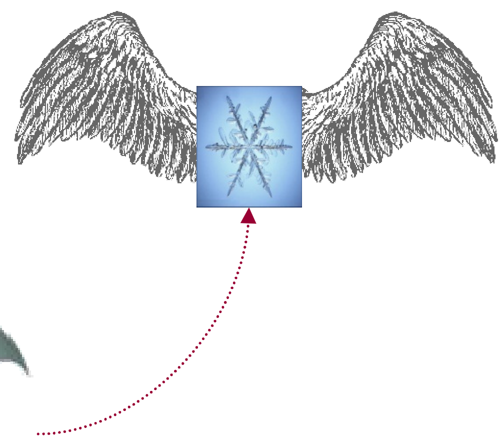
The free-energy principle
- Perception and generative models
- Hierarchies and predictive coding

Perception
- Birdsong and categorization
- Simulated lesions and perceptual uncertainty

Action
- Cued reaching and affordance
- Simulated lesions and behavioral uncertainty

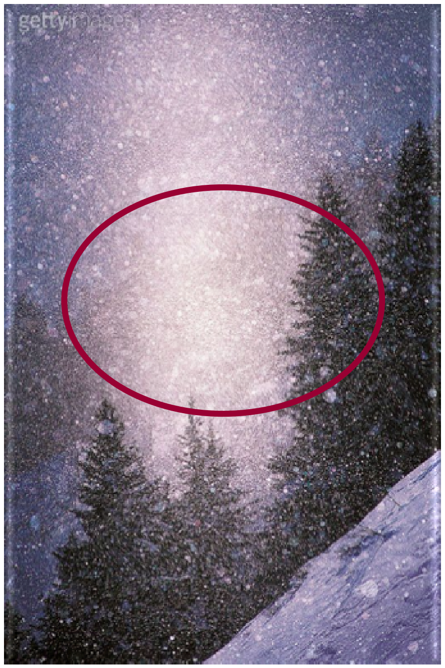
What is the difference between a snowflake and a bird?





…a bird can move (to avoid surprises)

What is the difference between snowfall and a flock of birds?



Ensemble dynamics, clumping and swarming

…birds (biological agents) stay in the same place

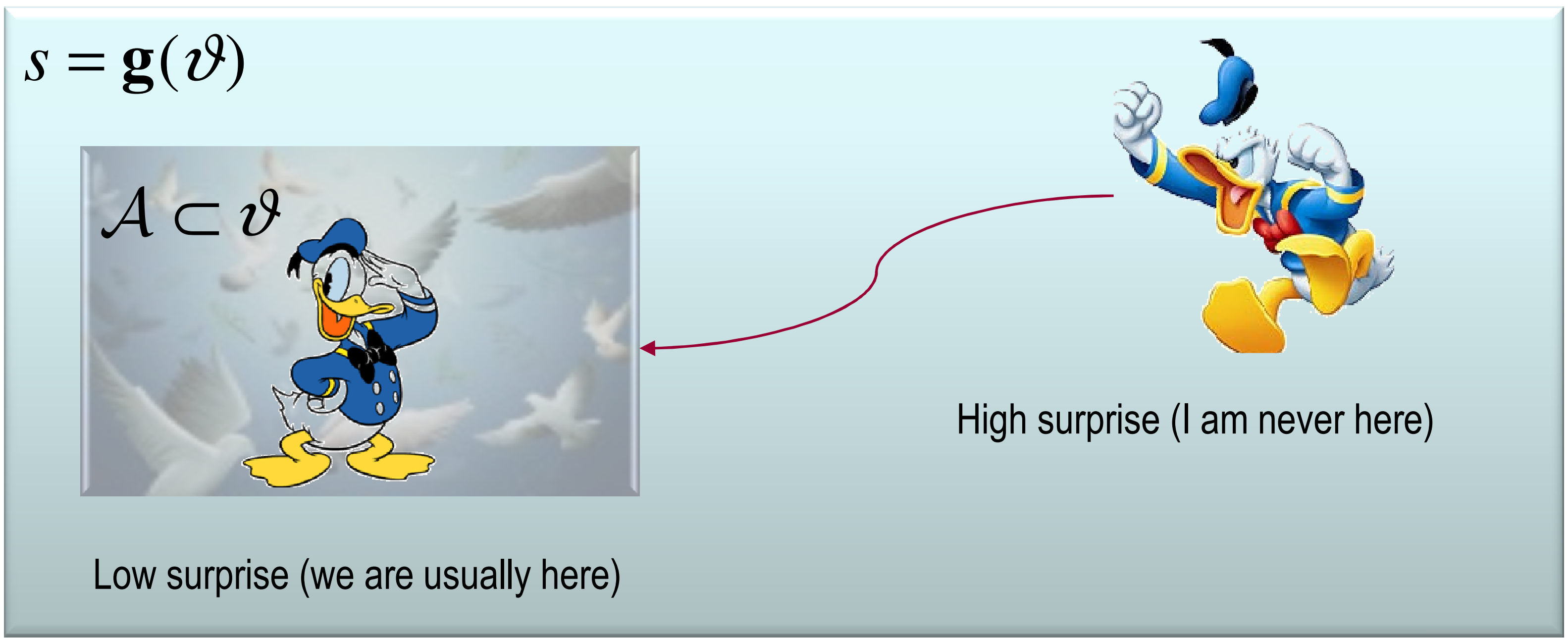
They resist the second law of thermodynamics, which says that their entropy should increase

But what is the entropy?

...entropy is just average surprise

$$H = \int_0^T dt \mathcal{L}(t) = -\int p(\vartheta \mid m) \ln p(\vartheta \mid m) d\vartheta$$

$$\mathcal{L} = -\ln p(s \mid m)$$



This means biological agents self-organize to minimise surprise. In other words, to ensure they occupy a limited number of states (cf homeostasis).

But there is a small problem… agents cannot measure their surprise

$$s = \mathbf{g}(\vartheta)$$

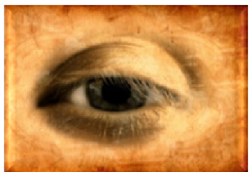
But they can measure their free-energy, which is always bigger than surprise

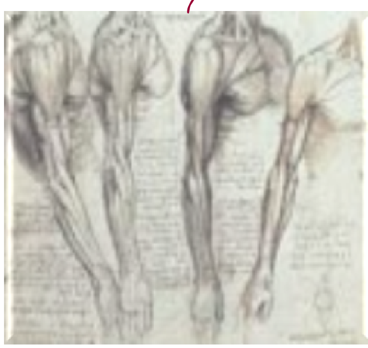
$$F(t) \geq \mathcal{L}(t)$$

This means agents should minimize their free-energy. So what is free-energy?

What is free-energy?

…free-energy is basically prediction error

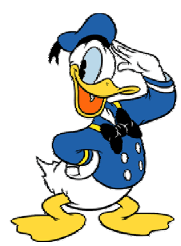




sensations – predictions

= prediction error

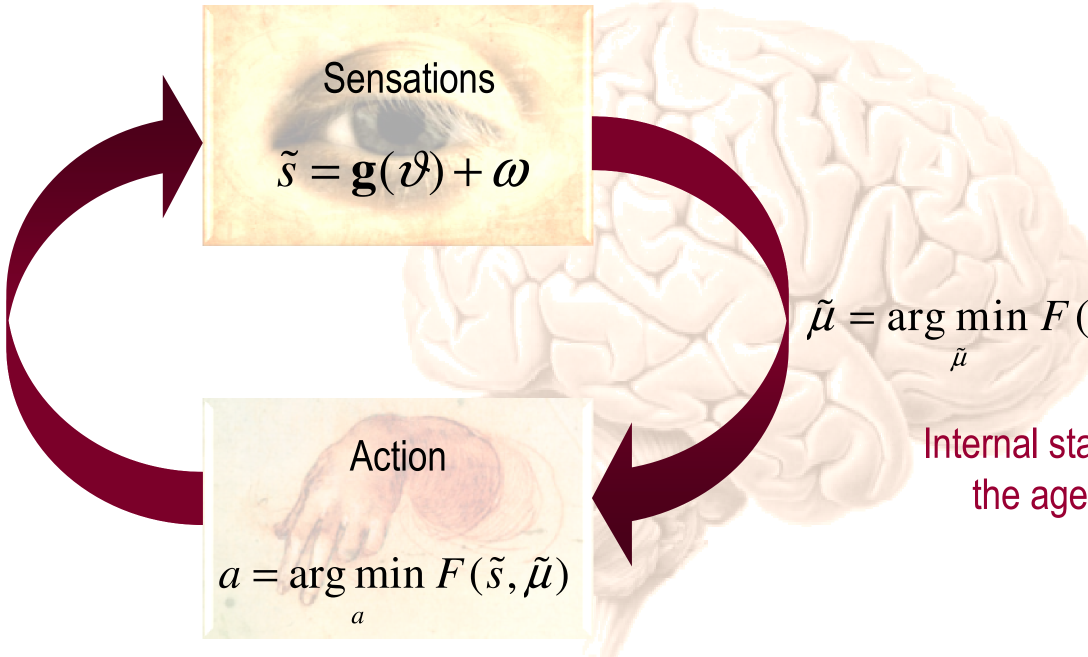
where small errors mean low surprise



More formally,

$$\dot{\vartheta} = \mathbf{f}(\vartheta, a) + \omega$$

External states
in the world



Sensations

$$\tilde{s} = \mathbf{g}(\vartheta) + \omega$$

Action

$$a = \arg\min_{a} F(\tilde{s}, \tilde{\mu})$$

$$\tilde{\mu} = \arg\min_{\tilde{\mu}} F(\tilde{s}, \tilde{\mu})$$

Internal states of
the agent (m)

Free-energy is a function of sensations and a proposal density over hidden causes

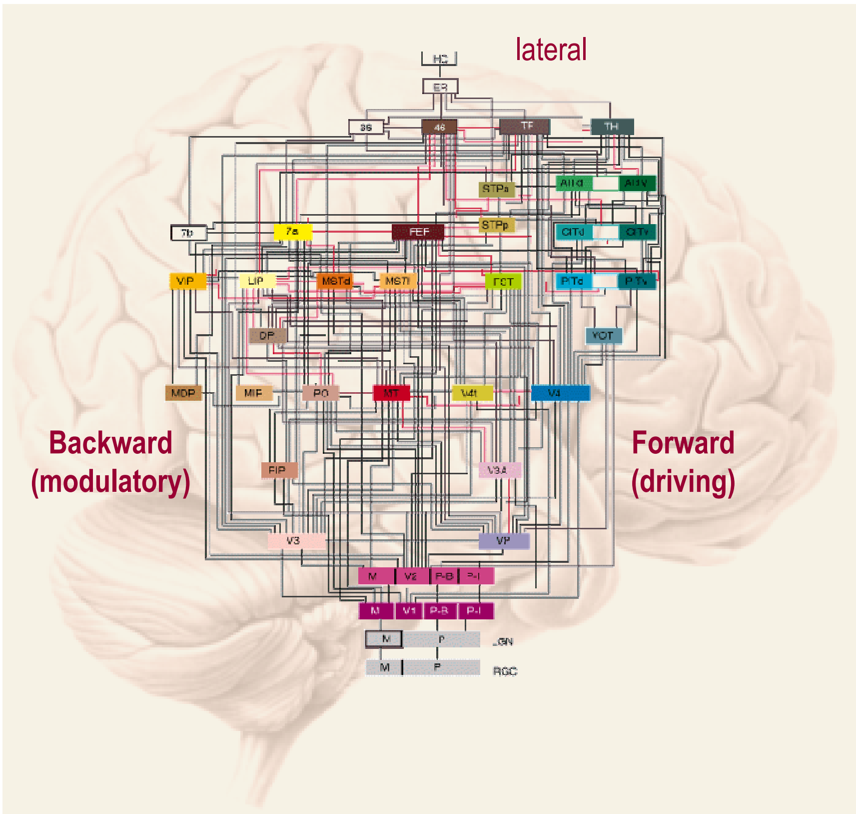
$$F(\tilde{s}, \tilde{\mu}) = Energy - Entropy = \mathbf{E}_q(G) + \mathbf{E}_q(\ln q(\vartheta \mid \tilde{\mu}))$$

and can be evaluated, given a generative model (Gibbs Energy) or likelihood and prior:

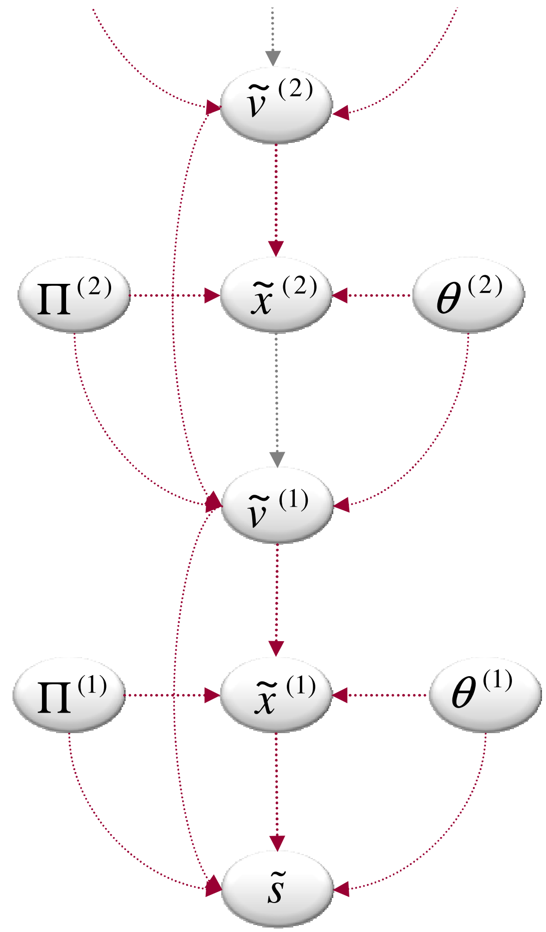
$$G(\tilde{s}, \vartheta) = -\ln p(\tilde{s}, \vartheta \mid m) = -\ln p(\tilde{s} \mid \vartheta, m) - \ln p(\vartheta \mid m)$$

So what models might the brain use?

Hierarchal models in the brain

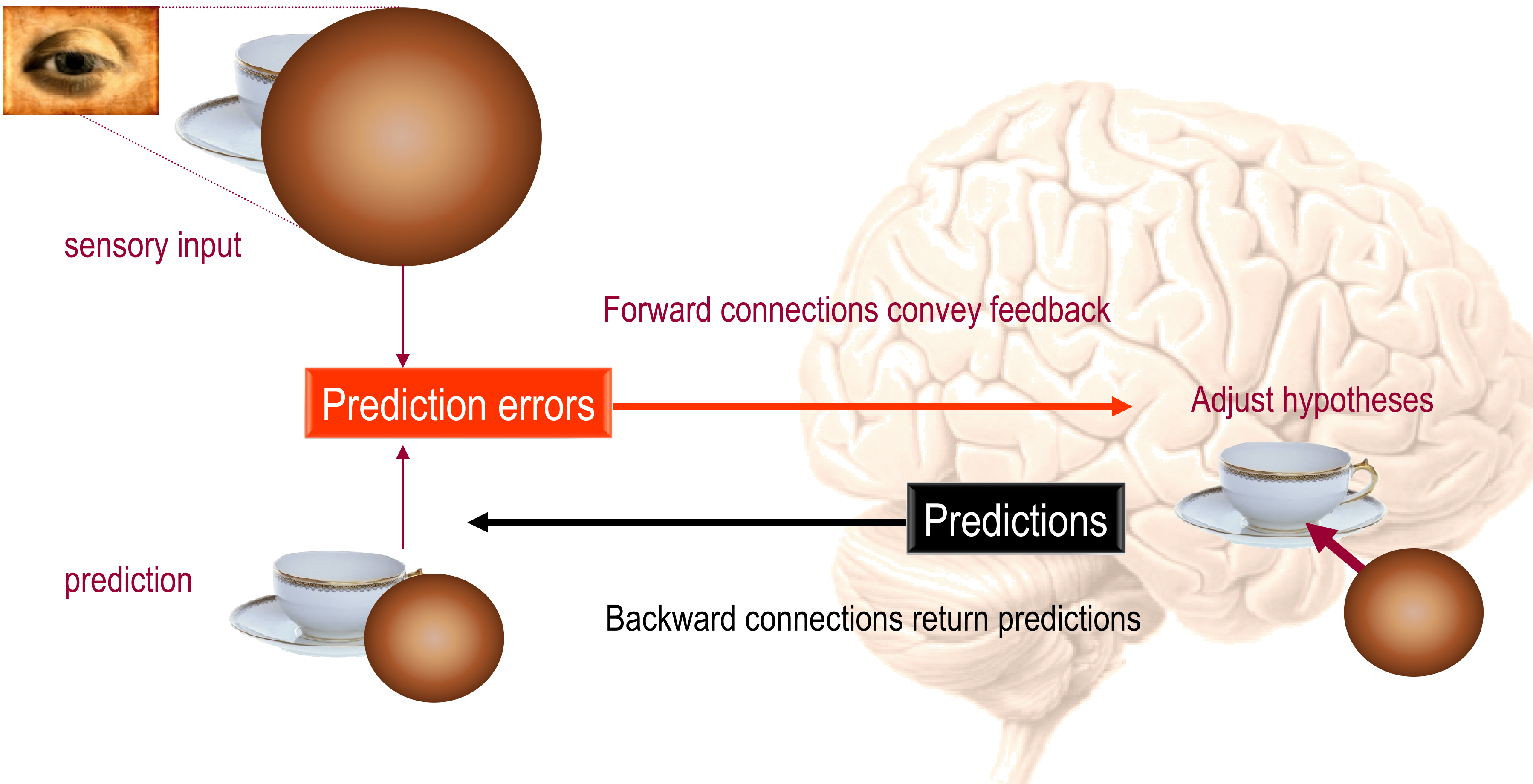


$$\tilde{v}^{(i-1)} = g(\tilde{x}^{(i)}, \tilde{v}^{(i)}, \theta^{(i)}) + \tilde{\omega}^{(v,i)}$$
$$\mathcal{D}\tilde{x}^{(i)} = f(\tilde{x}^{(i)}, \tilde{v}^{(i)}, \theta^{(i)}) + \tilde{\omega}^{(x,i)}$$



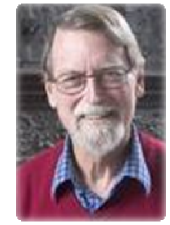
$$\vartheta \supset \{\tilde{x}(t), \tilde{v}(t), \theta, \Pi\}$$

So how do prediction errors change predictions?



...by hierarchical message passing in the brain

David Mumford

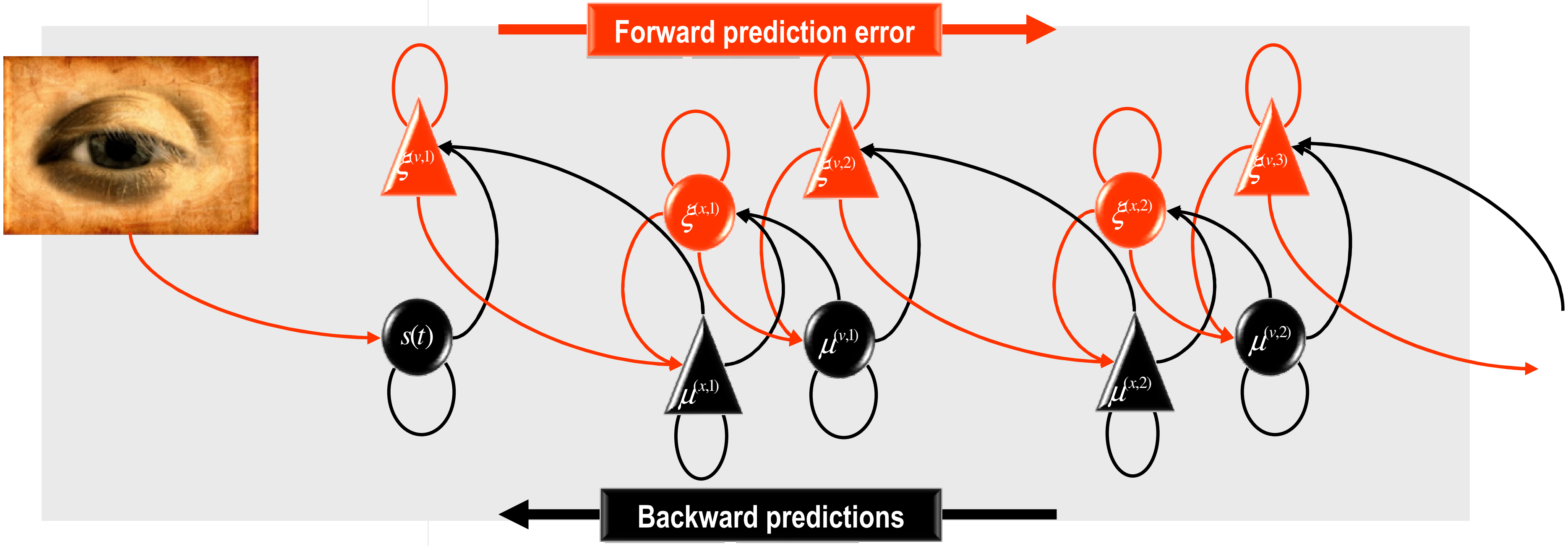


More formally,

Synaptic activity and message-passing

$$\xi^{(v,i)} = \Pi^{(v,i)}\tilde{\varepsilon}^{(v,i)} = \Pi^{(v,i)}(\tilde{\mu}^{(v,i-1)} - \tilde{g}^{(i)})$$

$$\xi^{(x,i)} = \Pi^{(x,i)}\tilde{\varepsilon}^{(x,i)} = \Pi^{(x,i)}(\mathcal{D}\tilde{\mu}^{(x,i)} - \tilde{f}^{(i)})$$



$$\dot{\tilde{\mu}}^{(v,i)} = \mathcal{D}\tilde{\mu}^{(v,i)} - \tilde{\varepsilon}_v^{(i)T}\xi^{(i)} - \xi^{(v,i+1)}$$

$$\dot{\tilde{\mu}}^{(x,i)} = \mathcal{D}\tilde{\mu}^{(x,i)} - \tilde{\varepsilon}_x^{(i)T}\xi^{(i)}$$

cf., Predictive coding or Kalman-Bucy filtering

Summary

Biological agents resist the second law of thermodynamics

They must minimize their average surprise (entropy)

They minimize surprise by suppressing prediction error (free-energy)

Prediction error can be reduced by changing predictions (perception)

Prediction error can be reduced by changing sensations (action)

Perception entails recurrent message passing in the brain to optimise predictions

Predictions depend upon the precision of prediction errors

Overview

Ensemble dynamics	Entropy and equilibria Free-energy and surprise
The free-energy principle	Perception and generative models Hierarchies and predictive coding

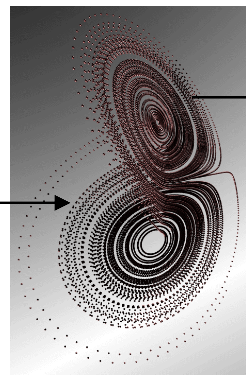
Perception	Birdsong and categorization Simulated lesions and perceptual uncertainty
Action	Cued reaching and affordance Simulated lesions and behavioral uncertainty

Making bird songs with Lorenz attractors

$$v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

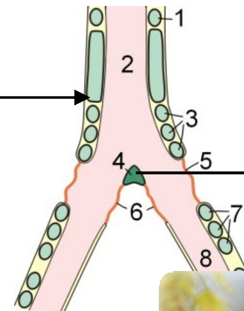
causal states

Vocal centre

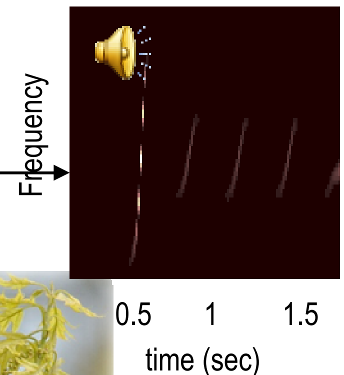


hidden states

Syrinx

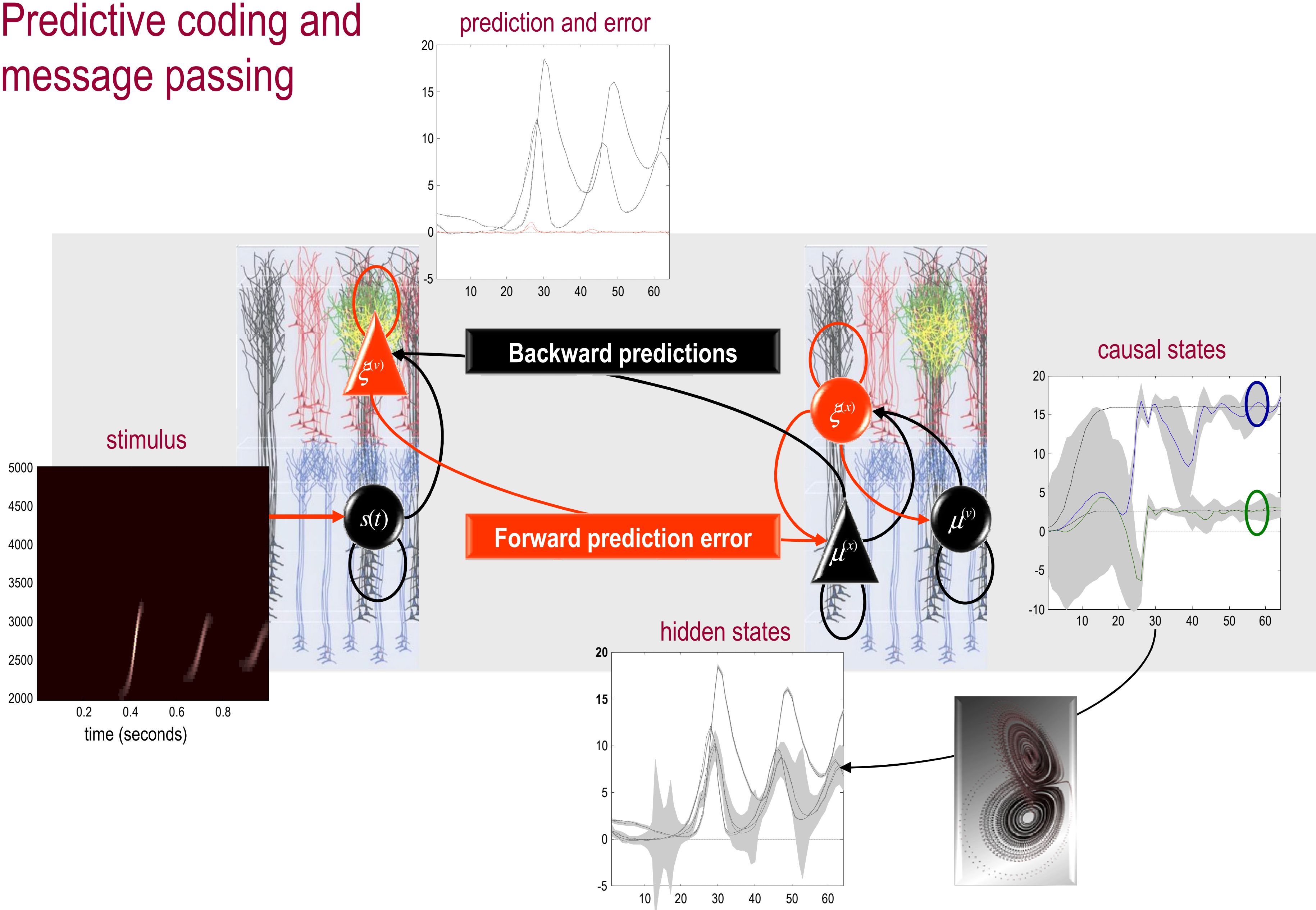


Sonogram

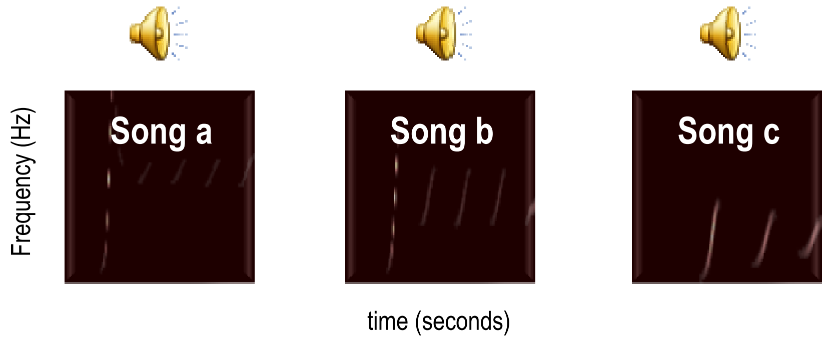


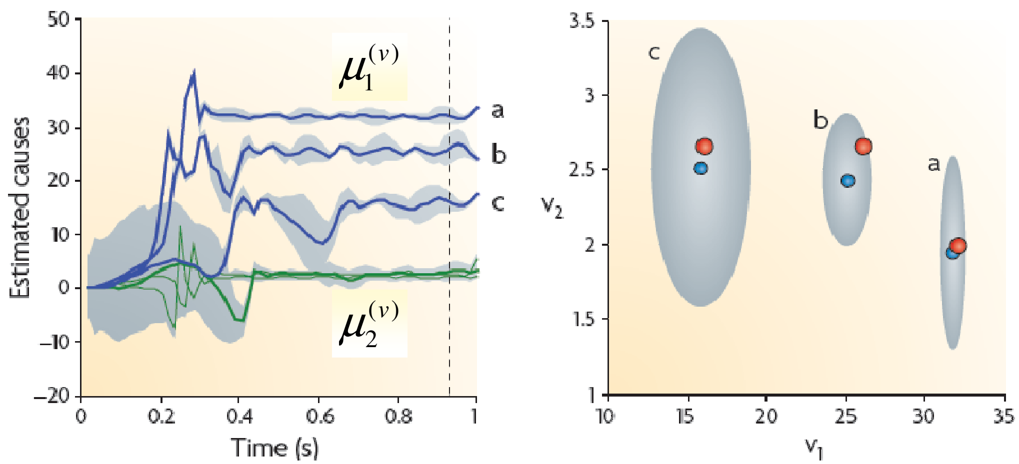
$$f^{(1)} = \begin{bmatrix} 18x_2^{(1)} - 18x_1^{(1)} \\ v_1^{(1)}x_1^{(1)} - 2x_3^{(1)}x_1^{(1)} - x_2^{(1)} \\ 2x_1^{(1)}x_2^{(1)} - v_2^{(1)}x_3^{(1)} \end{bmatrix}$$

Predictive coding and message passing



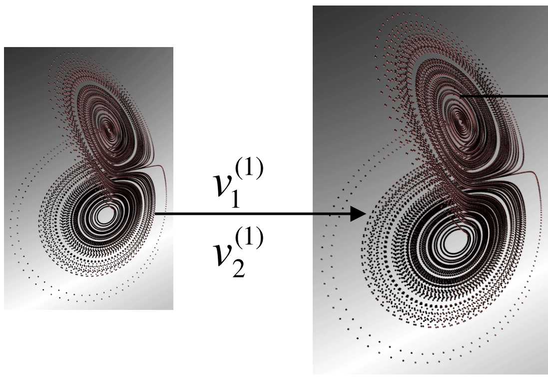
Perceptual categorization



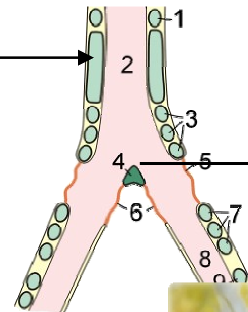


Hierarchical (itinerant) birdsong: sequences of sequences

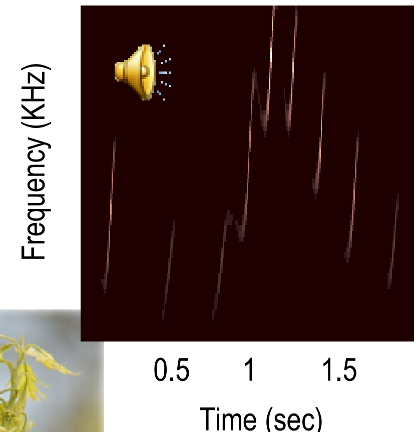
Neuronal hierarchy



Syrinx



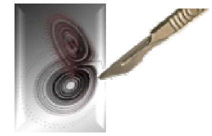
sonogram

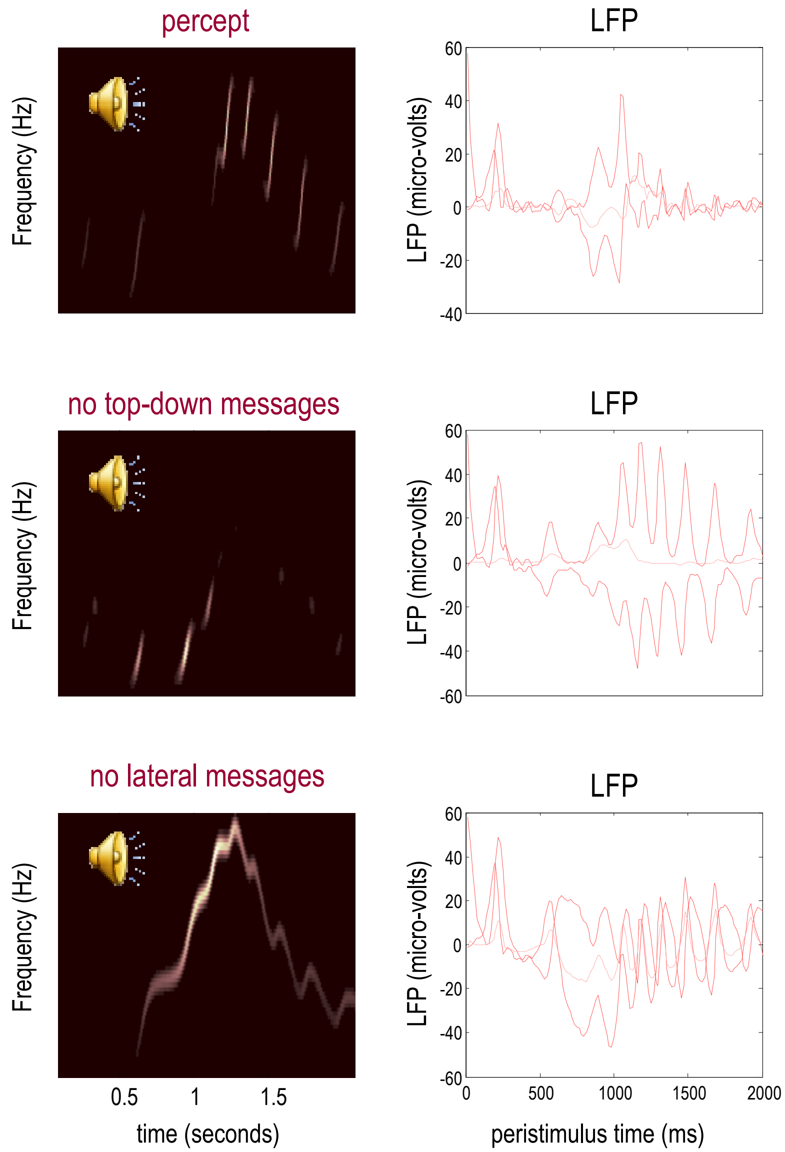


$$f^{(2)} = \begin{bmatrix} 18x_2^{(2)} - 18x_1^{(2)} \\ 32x_1^{(2)} - 2x_3^{(2)}x_1^{(2)} - x_2^{(2)} \\ 2x_1^{(2)}x_2^{(2)} - \frac{8}{3}x_3^{(2)} \end{bmatrix} \qquad f^{(1)} = \begin{bmatrix} 18x_2^{(1)} - 18x_1^{(1)} \\ v_1^{(1)}x_1^{(1)} - 2x_3^{(1)}x_1^{(1)} - x_2^{(1)} \\ 2x_1^{(1)}x_2^{(1)} - v_2^{(1)}x_3^{(1)} \end{bmatrix}$$

$$g^{(2)} = \begin{bmatrix} x_2^{(2)} \\ x_3^{(2)} \end{bmatrix} = \begin{bmatrix} v_1^{(1)} \\ v_2^{(1)} \end{bmatrix} \qquad g^{(1)} = \begin{bmatrix} x_2^{(1)} \\ x_3^{(1)} \end{bmatrix} = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$$

Simulated lesions and false inference





$$\dot{\tilde{\mu}}^{(v,i)} = \mathcal{D}\tilde{\mu}^{(v,i)} - \tilde{\varepsilon}_v^{(i)T}\xi^{(i)} - \xi^{(v,i+1)}$$
$$\dot{\tilde{\mu}}^{(x,i)} = \mathcal{D}\tilde{\mu}^{(x,i)} - \tilde{\varepsilon}_x^{(i)T}\xi^{(i)}$$

no structural priors

$$\dot{\tilde{\mu}}^{(v,i)} = \mathcal{D}\tilde{\mu}^{(v,i)} - \tilde{\varepsilon}_v^{(i)T}\xi^{(i)} - \cancel{\xi^{(v,i+1)}}$$
$$\dot{\tilde{\mu}}^{(x,i)} = \mathcal{D}\tilde{\mu}^{(x,i)} - \tilde{\varepsilon}_x^{(i)T}\xi^{(i)}$$

no dynamical priors

$$\dot{\tilde{\mu}}^{(v,i)} = \cancel{\mathcal{D}\tilde{\mu}^{(v,i)}} - \tilde{\varepsilon}_v^{(i)T}\xi^{(i)} - \xi^{(v,i+1)}$$
$$\dot{\tilde{\mu}}^{(x,i)} = \cancel{\mathcal{D}\tilde{\mu}^{(x,i)}} - \tilde{\varepsilon}_x^{(i)T}\xi^{(i)}$$

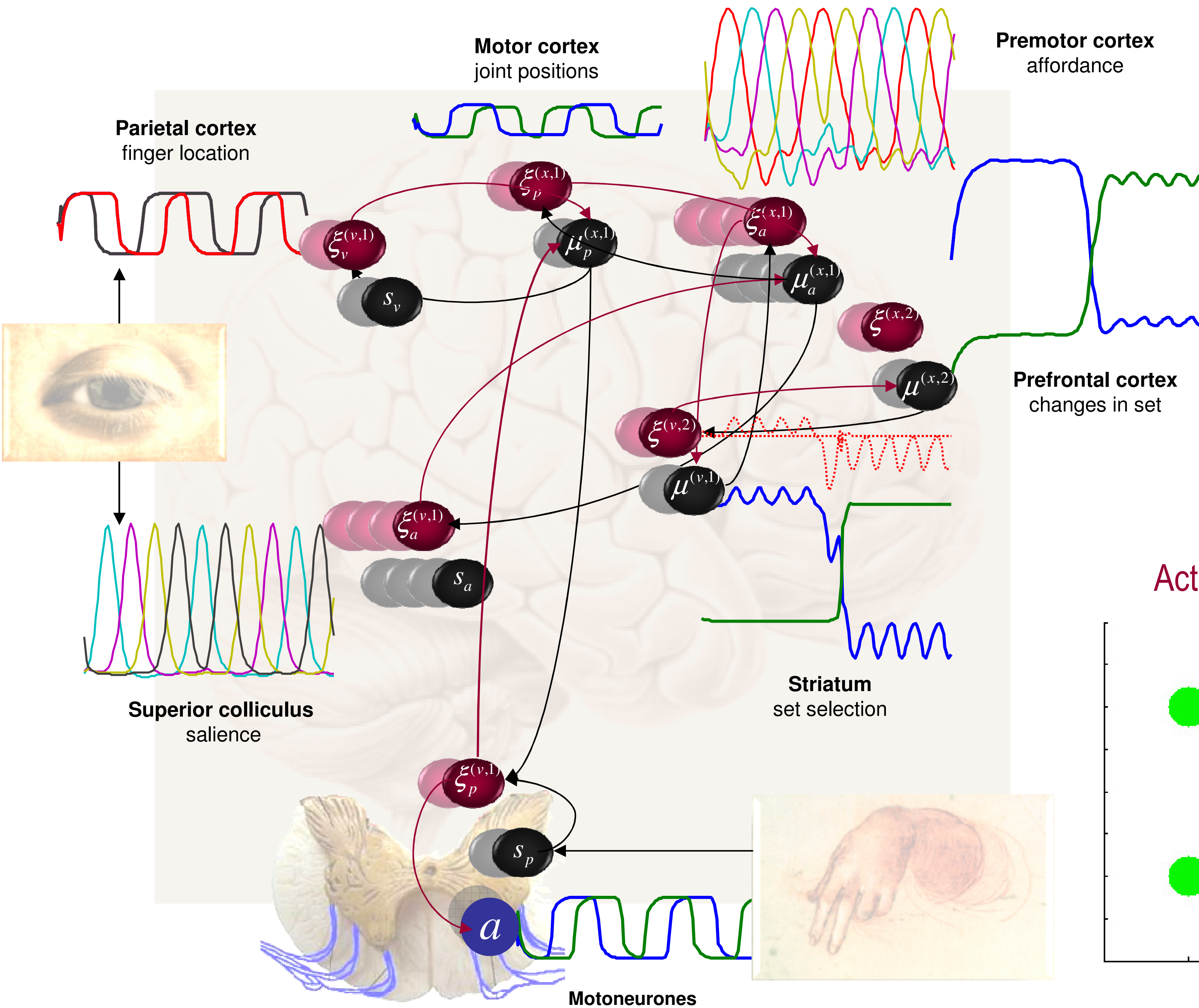
Overview

Ensemble dynamics
- Entropy and equilibria
- Free-energy and surprise

The free-energy principle
- Perception and generative models
- Hierarchies and predictive coding

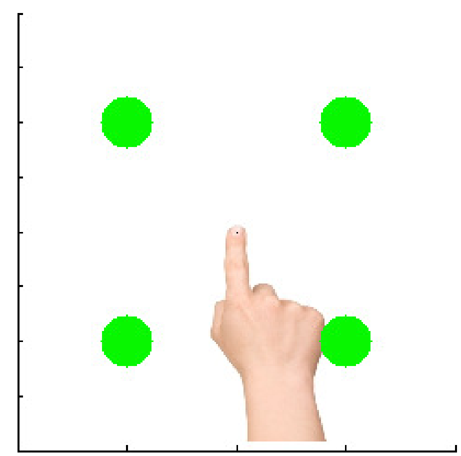
Perception
- Birdsong and categorization
- Simulated lesions and perceptual uncertainty

Action
- Cued reaching and affordance
- Simulated lesions and behavioral uncertainty



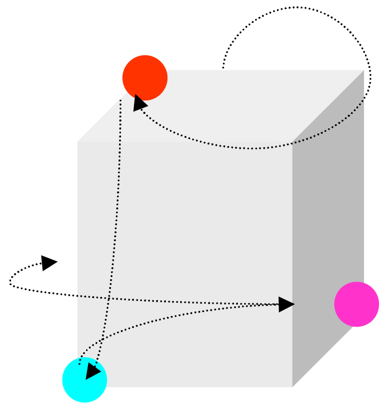
Anatol Feldman

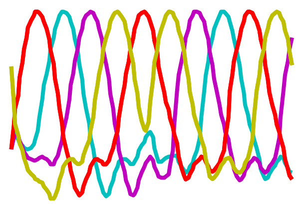
Active inference



Misha Rabinovich

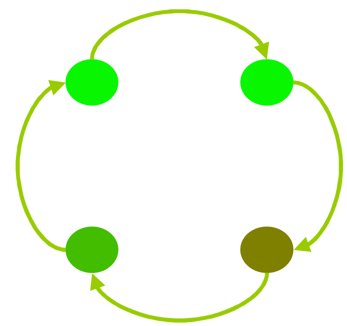
Lotka-Volterra dynamics: winnerless competition





$$\xrightarrow{\ell \cdot s(x)}$$

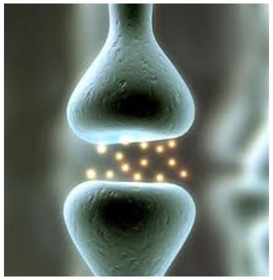
$$\ell = \begin{bmatrix} -1 & 1 & 1 & -1 \\ -1 & -1 & 1 & 1 \end{bmatrix}$$



$$\sigma(x, v) = \mathbf{A}(v)(\vec{1} + e^{x})^{-1} - \tfrac{1}{8}x + \vec{1}$$

$$s(x) = \frac{e^{x}}{e^{xT}\vec{1}}$$

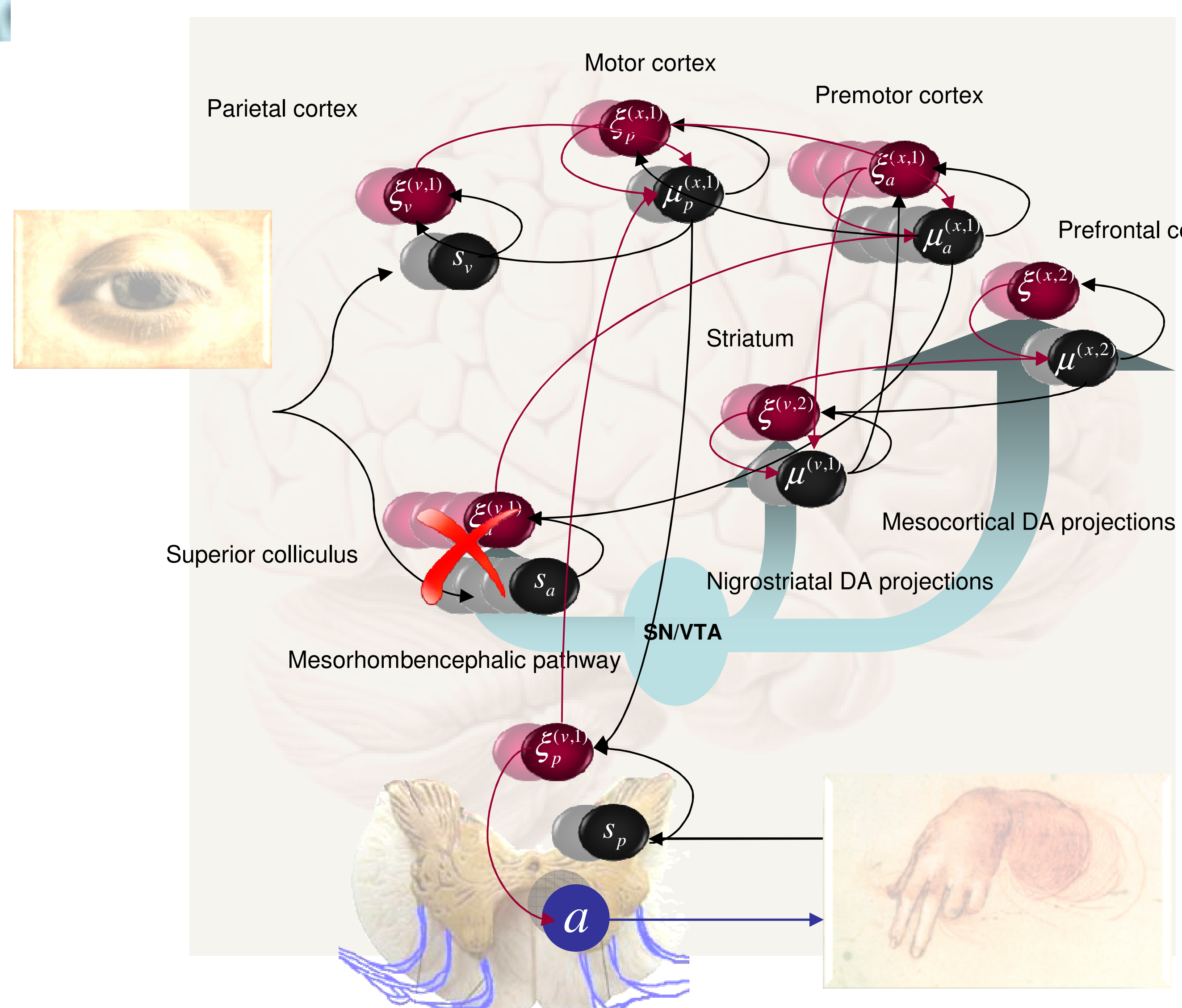
$$\mathbf{A}(v) = \begin{bmatrix} 0 & -v & 0 & v \\ v & 0 & -v & 0 \\ 0 & v & 0 & -v \\ -v & 0 & v & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

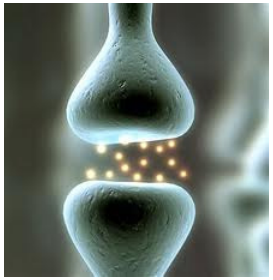


Dopamine and precision

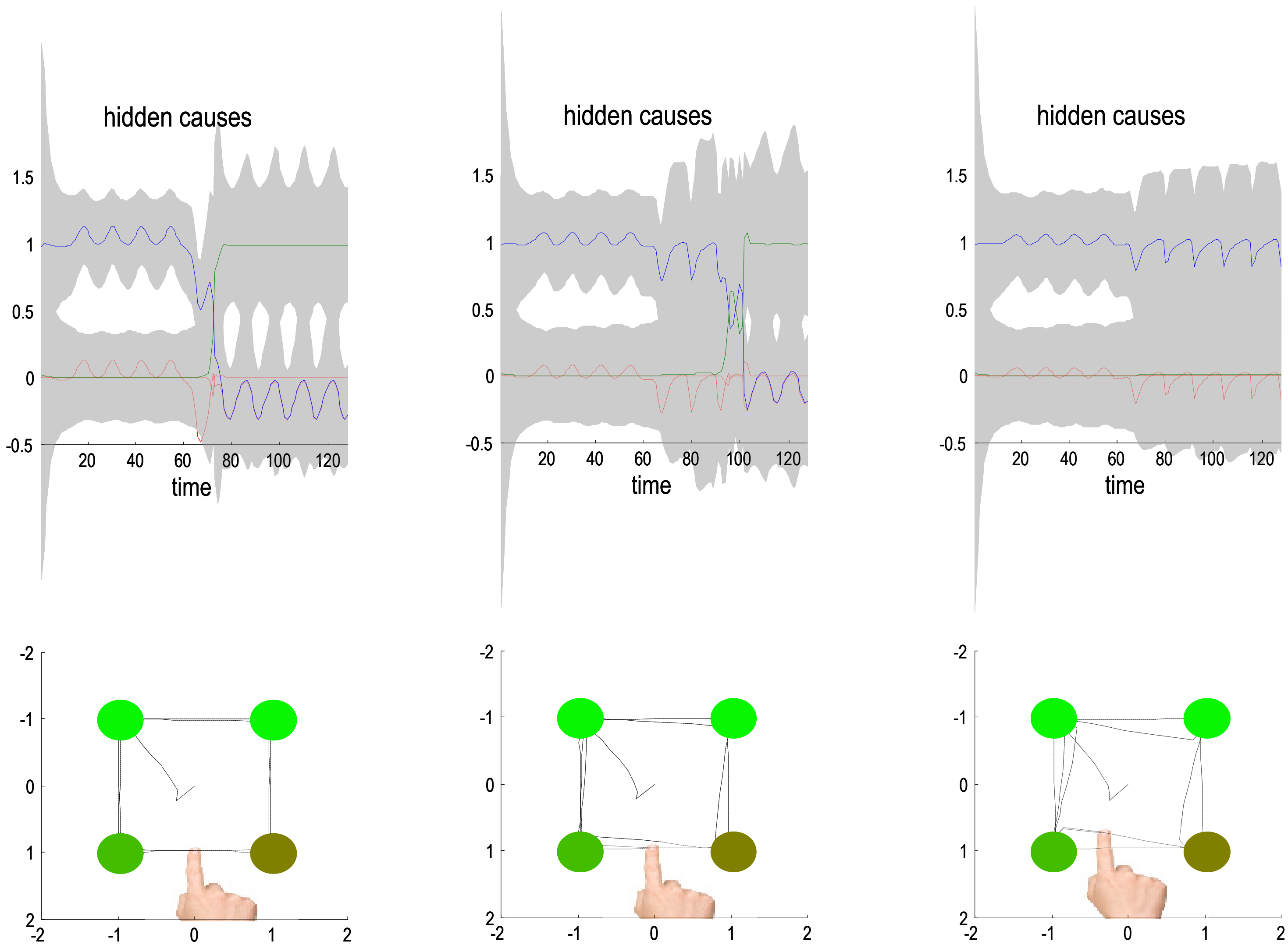
$$\xi^{(v,i)} = \Pi^{(v,i)}\tilde{\varepsilon}^{(v,i)} = \Pi^{(v,i)}(\tilde{\mu}^{(v,i-1)} - \tilde{g}^{(i)})$$

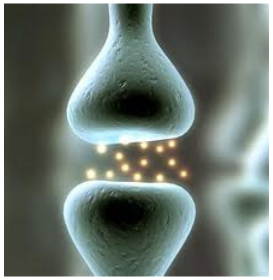
$$\xi^{(x,i)} = \Pi^{(x,i)}\tilde{\varepsilon}^{(x,i)} = \Pi^{(x,i)}(\mathcal{D}\tilde{\mu}^{(x,i)} - \tilde{f}^{(i)})$$



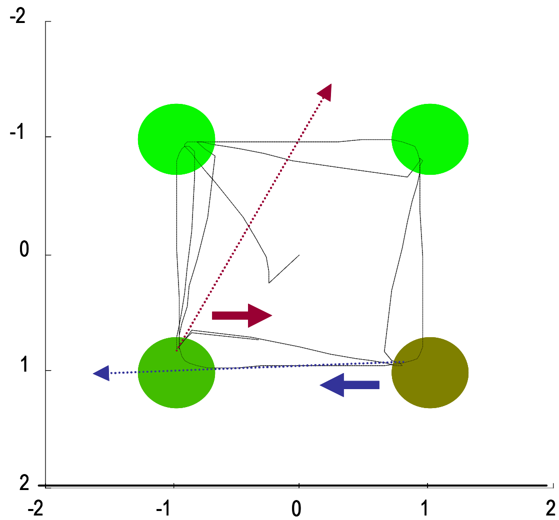


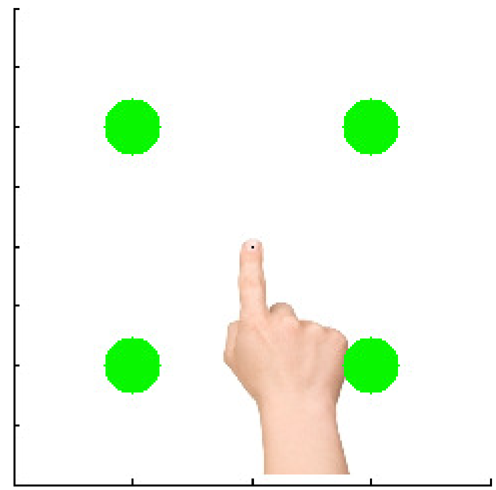
$$\tilde{\Pi}_a^{(1,v)} = \exp(5.0) \qquad \tilde{\Pi}_a^{(1,v)} = \exp(3.5) \qquad \tilde{\Pi}_a^{(1,v)} = \exp(2.5)$$

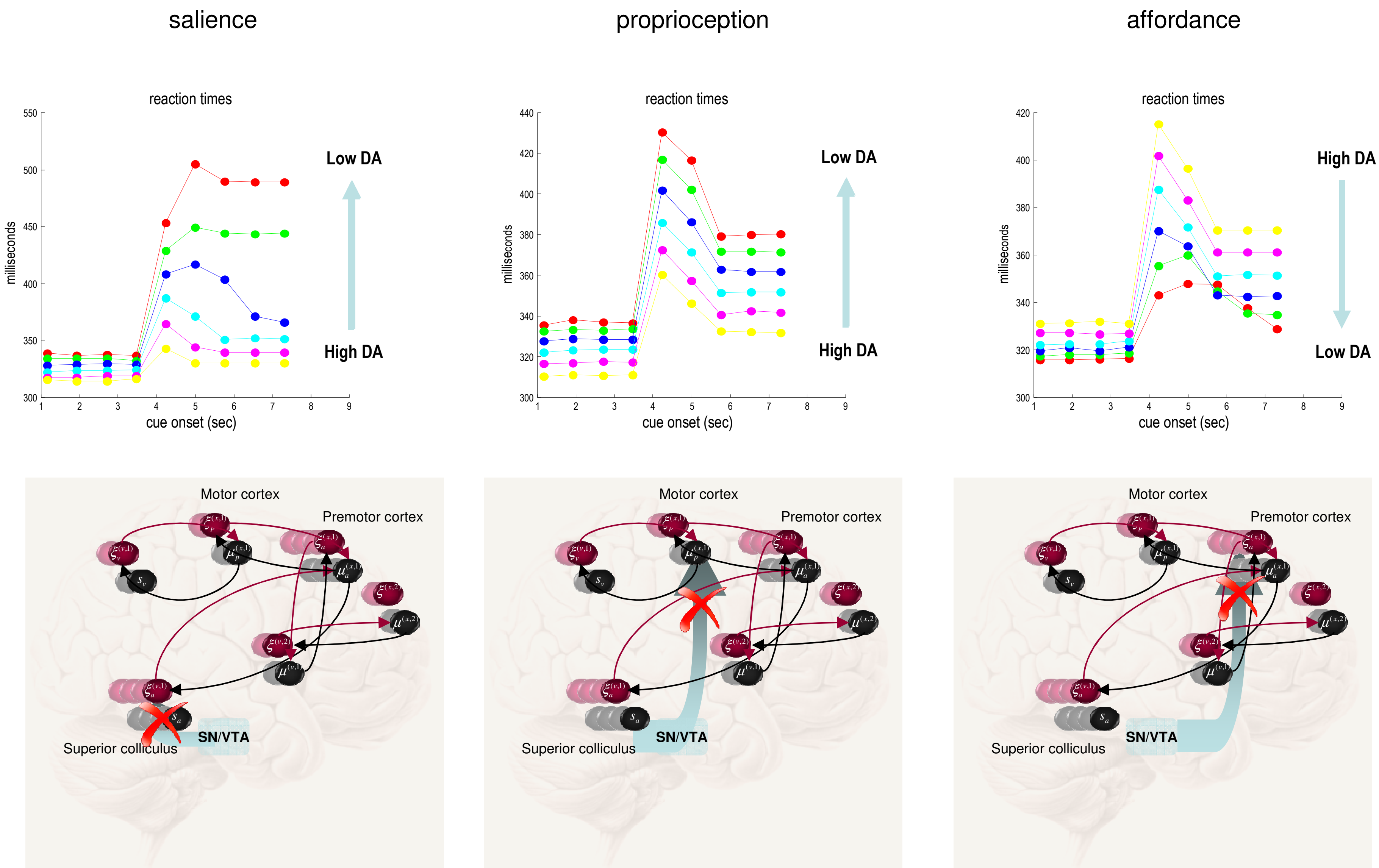




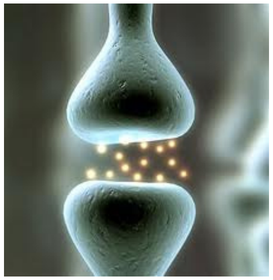
Uncertainty and perseveration



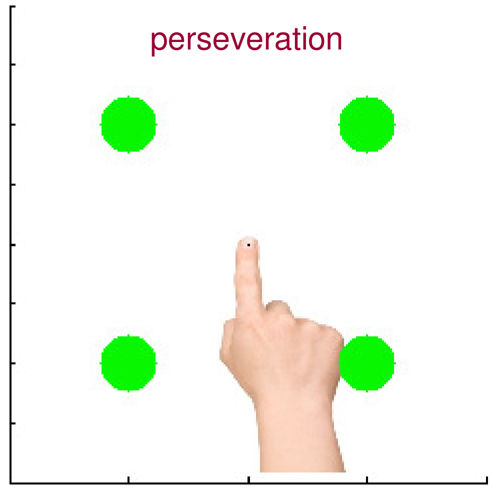


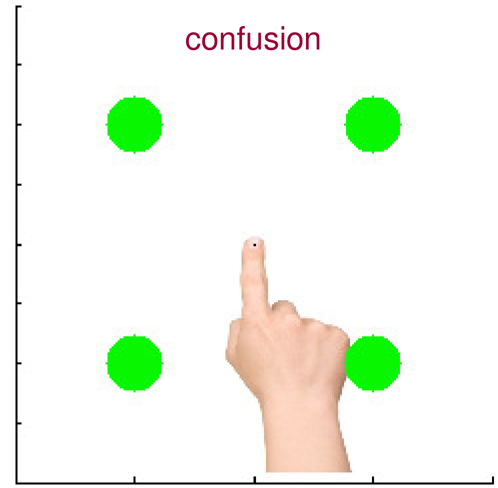


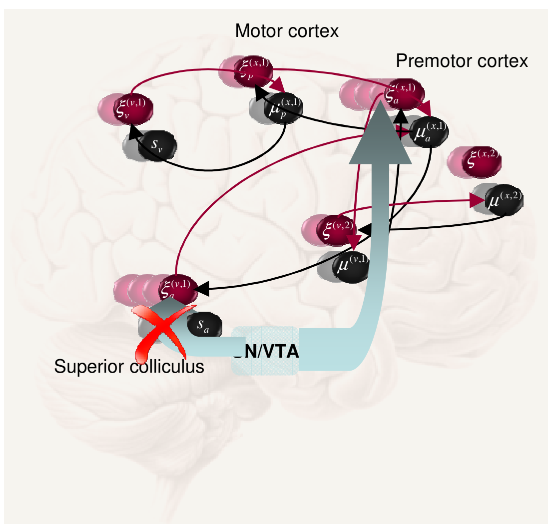
salience
proprioception
affordance
reaction times
milliseconds
cue onset (sec)
Low DA
High DA
Motor cortex
Premotor cortex
Superior colliculus
SN/VTA

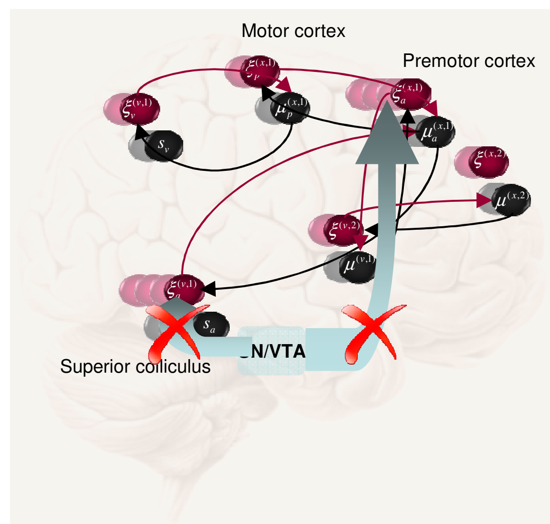


Uncertainly, delusions and confusion









Thank you

And thanks to collaborators:

Rick Adams
Harriet Brown
Jean Daunizeau
Lee Harrison
Stefan Kiebel
James Kilner
Jérémie Mattout
Klaas Stephan

And colleagues:

Peter Dayan
Jörn Diedrichsen
Paul Verschure
Florentin Wörgötter

And many others