Life as we know it

Karl Friston, University College London

How much about our interaction with – and experience of – our world can be deduced from basic principles? This talk reviews recent attempts to understand the self-organised behaviour of embodied agents – like ourselves – as satisfying basic imperatives for sustained exchanges with our world. In brief, one simple driving force appears to explain nearly every aspect of our behaviour and experience. This driving force is the minimisation of surprise or prediction error. In the context of perception, this corresponds to (Bayes-optimal) predictive coding that suppresses exteroceptive prediction errors. In the context of action, simple reflexes can be seen as suppressing proprioceptive prediction errors. We will look at some of the phenomena that emerge from this formulation, such as hierarchical message passing in the brain and the perceptual inference that ensues. I hope to illustrate these points using simple simulations of how life-like behaviour emerges almost inevitably from coupled dynamical systems – and how this behaviour can be understood in terms of perception, action and action observation.

Overview

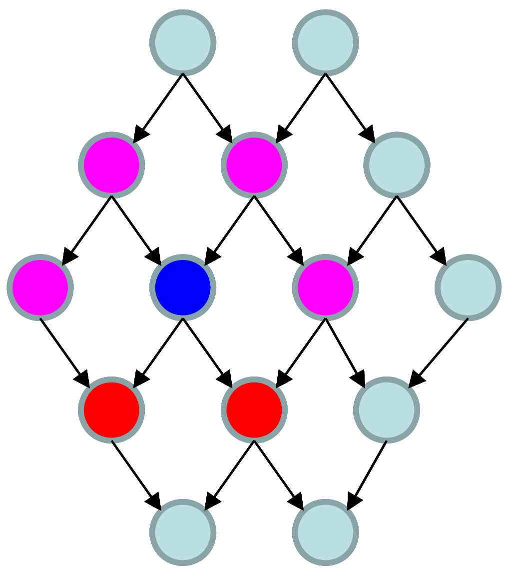
The statistics of life	Markov blankets and ergodic systems simulations of a primordial soup
The anatomy of inference	graphical models and predictive coding canonical microcircuits
Action and perception	perceptual omission responses simulations of action observation

"How can the events in space and time which take place within the spatial boundary of a living organism be accounted for by physics and chemistry?"

(Erwin Schrödinger 1943)

The Markov blanket as a statistical boundary

(parents, children and parents of children)



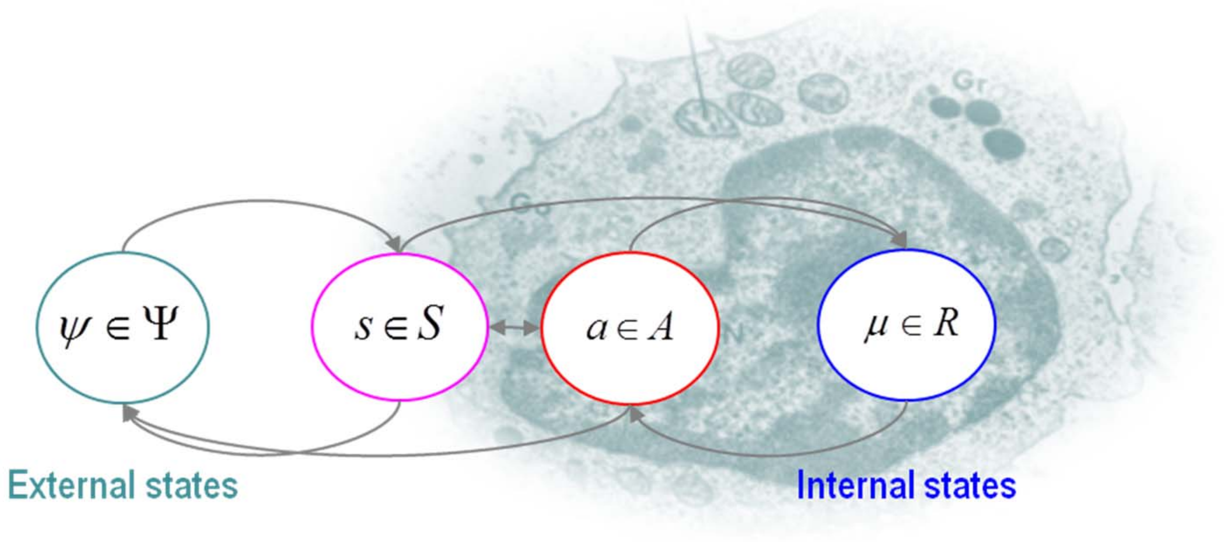
Internal states

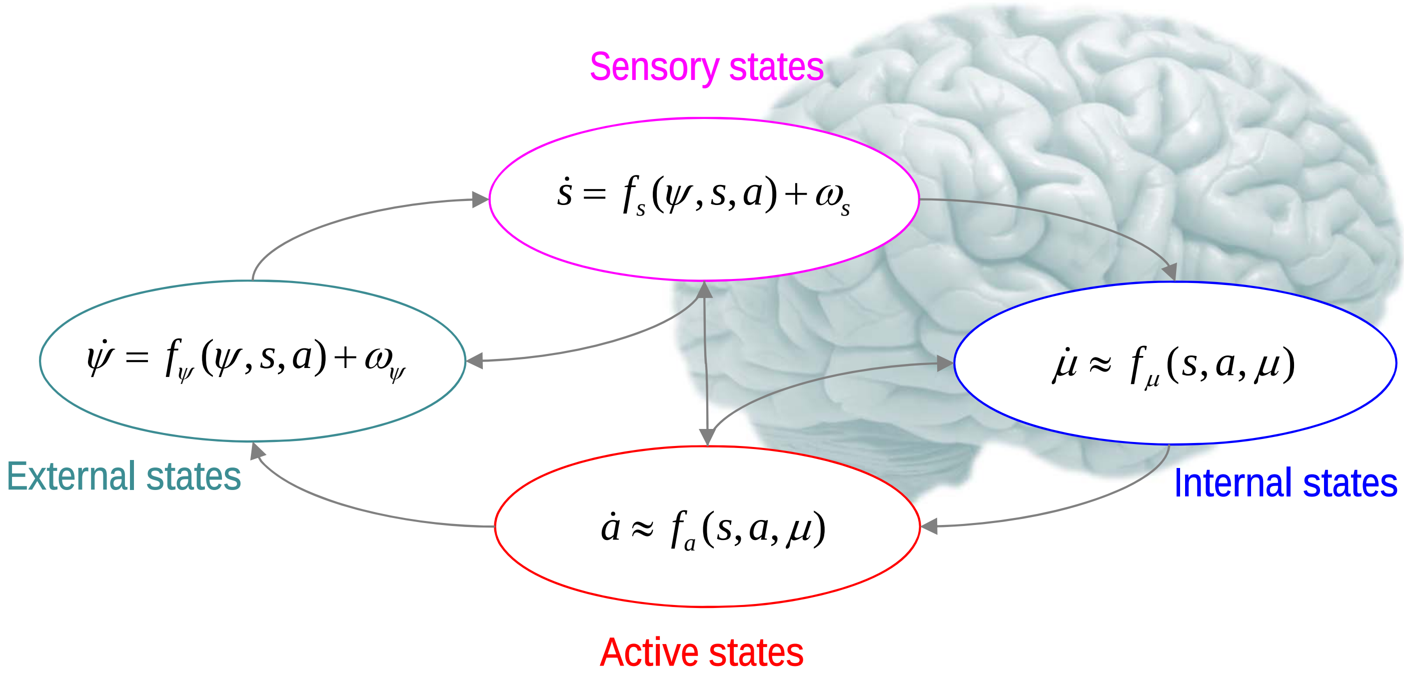
External states

Sensory states

Active states

The Markov blanket in biotic systems





lemma: *any (ergodic random) dynamical system (m) that possesses a Markov blanket will appear to actively maintain its structural and dynamical integrity*

$$\dot{x} = f(x) + \omega$$

$$p(x \mid m)$$

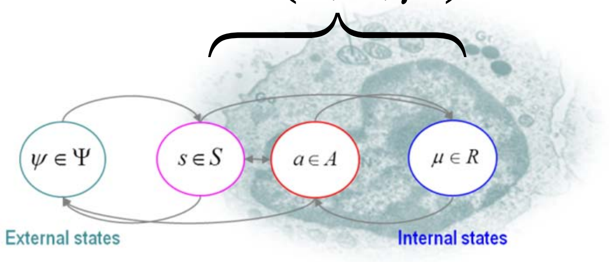
The Fokker-Planck equation $\dot{p}(x \mid m) = \nabla \cdot (\Gamma \nabla - f)p$

And its solution in terms of curl-free and divergence-free components

$$\dot{p}(x \mid m) = 0 \Leftrightarrow f(x) = (\Gamma - Q)\nabla \ln p(x \mid m)$$

But what about the Markov blanket?

$$\underbrace{\tilde{s}=(s,a,\mu)}$$



$$f_r(\tilde{s}) = (\Gamma - Q)\nabla_\mu \ln p(\tilde{s} \mid m)$$ Perception

$$f_a(\tilde{s}) = (\Gamma - Q)\nabla_a \ln p(\tilde{s} \mid m)$$ Action

$\ln p(\tilde{s} \mid m) =$ Value → Reinforcement learning, optimal control and utility theory

$-\ln p(\tilde{s} \mid m) =$ Surprise (free energy) → Information theory and minimum redundancy

$E_t[-\ln p(\tilde{s} \mid m)] =$ Entropy → Self-organisation, cybernetics, homoeostasis and autopoiesis

$p(\tilde{s} \mid m) =$ Model evidence → Bayesian brain, predictive coding and active inference

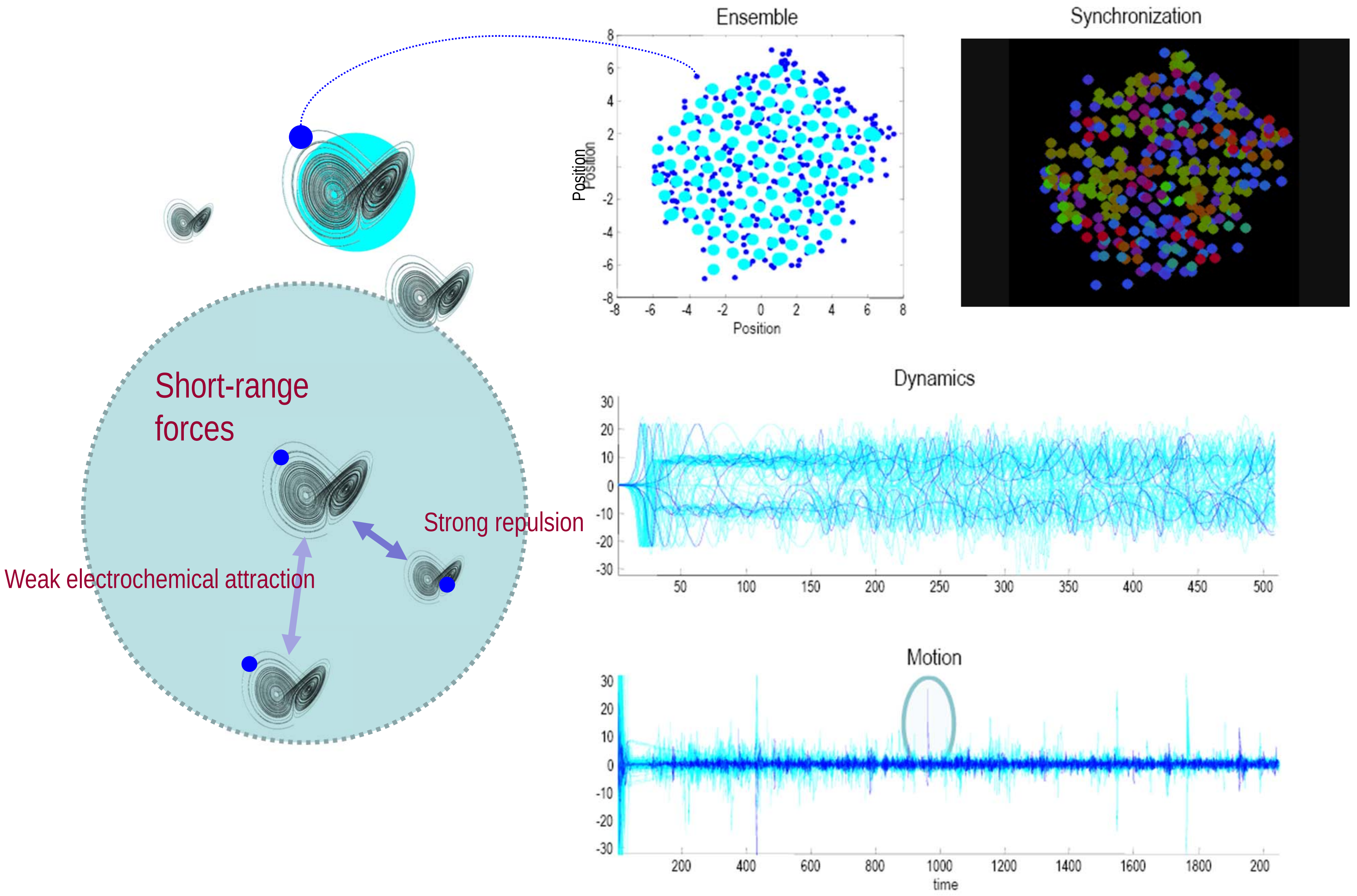
Overview

The statistics of life
- Markov blankets and ergodic systems
- simulations of a primordial soup

The anatomy of inference
- graphical models and predictive coding
- canonical microcircuits

Action and perception
- perceptual omission responses
- simulations of action observation

Simulations of a (prebiotic) primordial soup



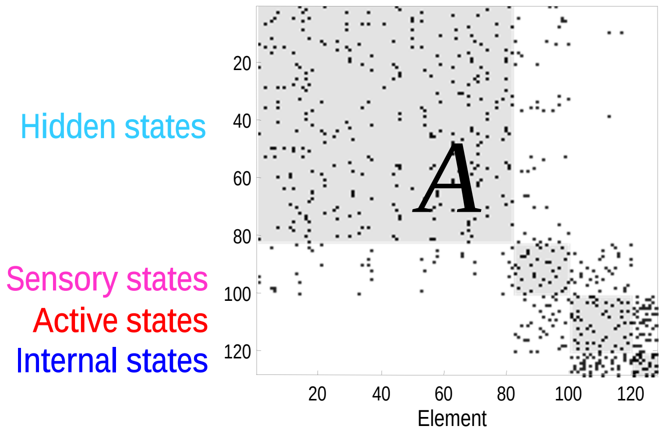
Finding the (principal) Markov blanket

Markov blanket matrix: encoding the children, parents and parents of children

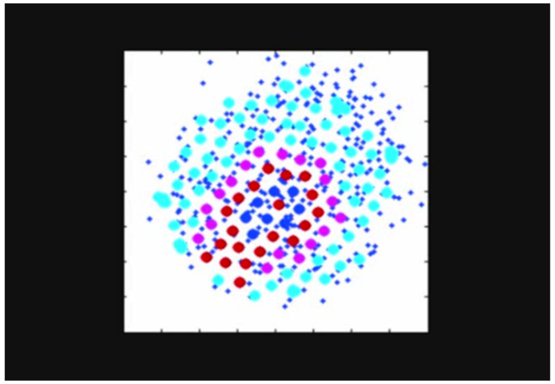
$$B = A + A^T + A^T A$$

Markov Blanket = $[B \cdot [\boldsymbol{eig(B) > \tau}]]$

Adjacency matrix

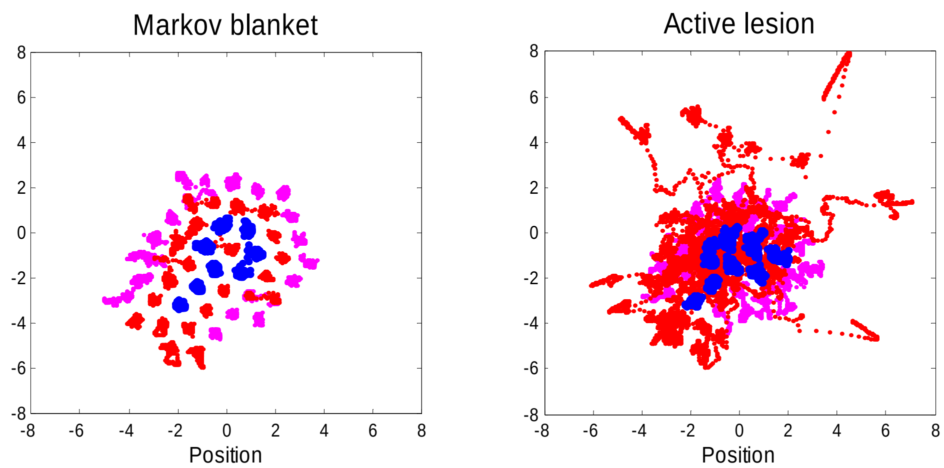


Markov Blanket

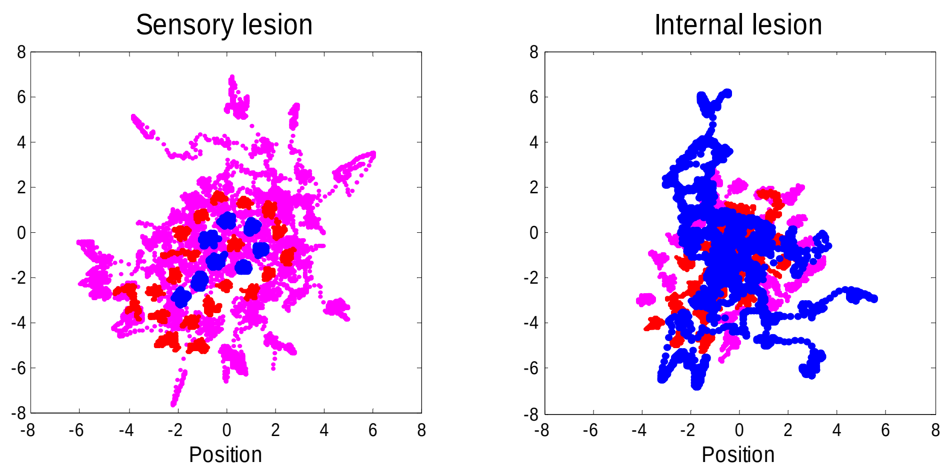


Does action maintain the structural and functional integrity of the Markov blanket (**autopoiesis**) ?

Do internal states appear to infer the hidden causes of sensory states (**active inference**) ?

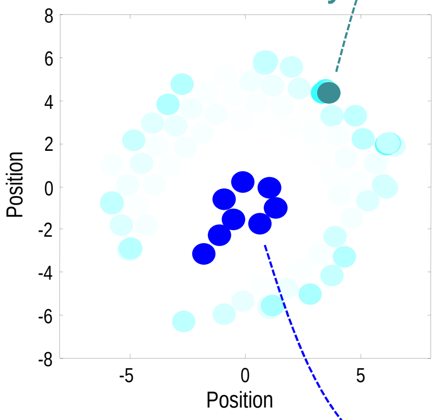


Autopoiesis, oscillator death and simulated brain lesions

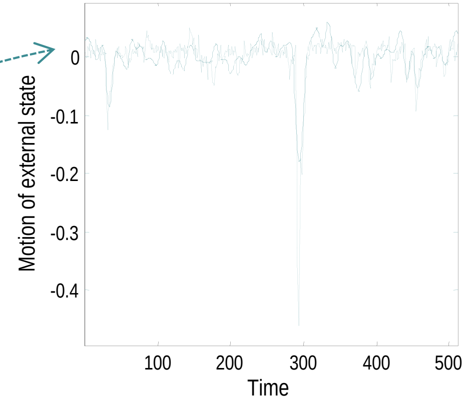


Decoding through the Markov blanket and simulated brain activation

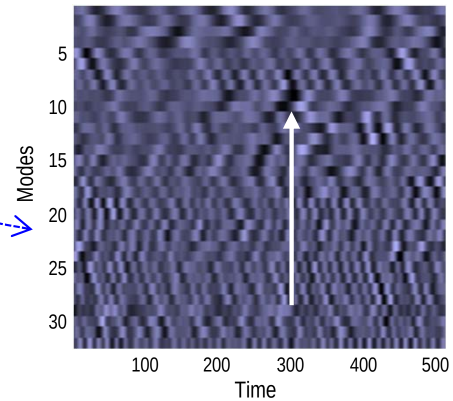
Predictability



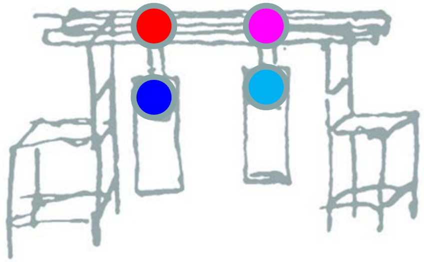
True and predicted motion



Internal states



Christiaan Huygens



The existence of a Markov blanket necessarily implies a partition of states into internal states, their Markov blanket (sensory and active states) and external or hidden states.

Because active states change – but are not changed by – external states they minimize the entropy of internal states and their Markov blanket. This means action will appear to maintain the structural and functional integrity of the Markov blanket (**autopoiesis**).

Internal states appear to infer the hidden causes of sensory states (by maximizing Bayesian evidence) and influence those causes though action (**active inference**)

Overview

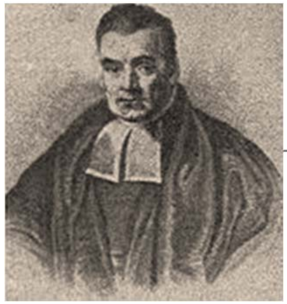
The statistics of life	Markov blankets and ergodic systems simulations of a primordial soup
The anatomy of inference	graphical models and predictive coding canonical microcircuits
Action and perception	perceptual omission responses simulations of action observation

Hermann von Helmholtz

"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" - von Helmholtz

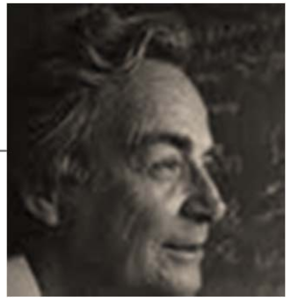
Richard Gregory

Geoffrey Hinton



Thomas Bayes

The Helmholtz machine and the Bayesian brain



Richard Feynman

“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” - von Helmholtz

Hermann von Helmholtz

Richard Gregory

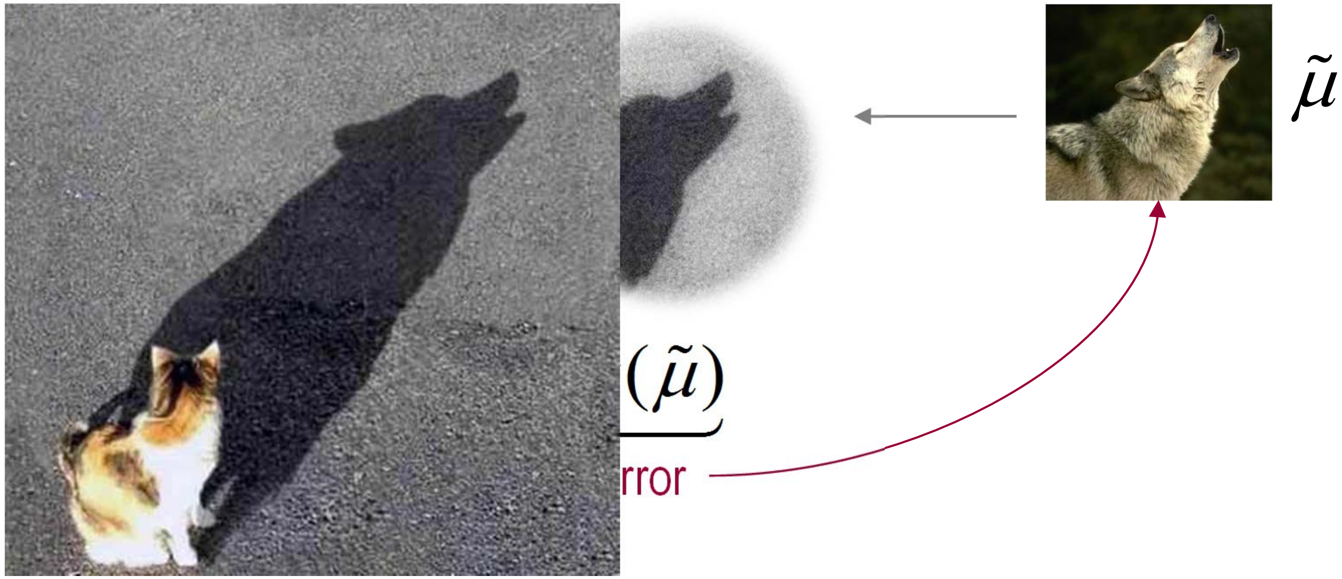
Impressions on the Markov blanket…

$s \in S$

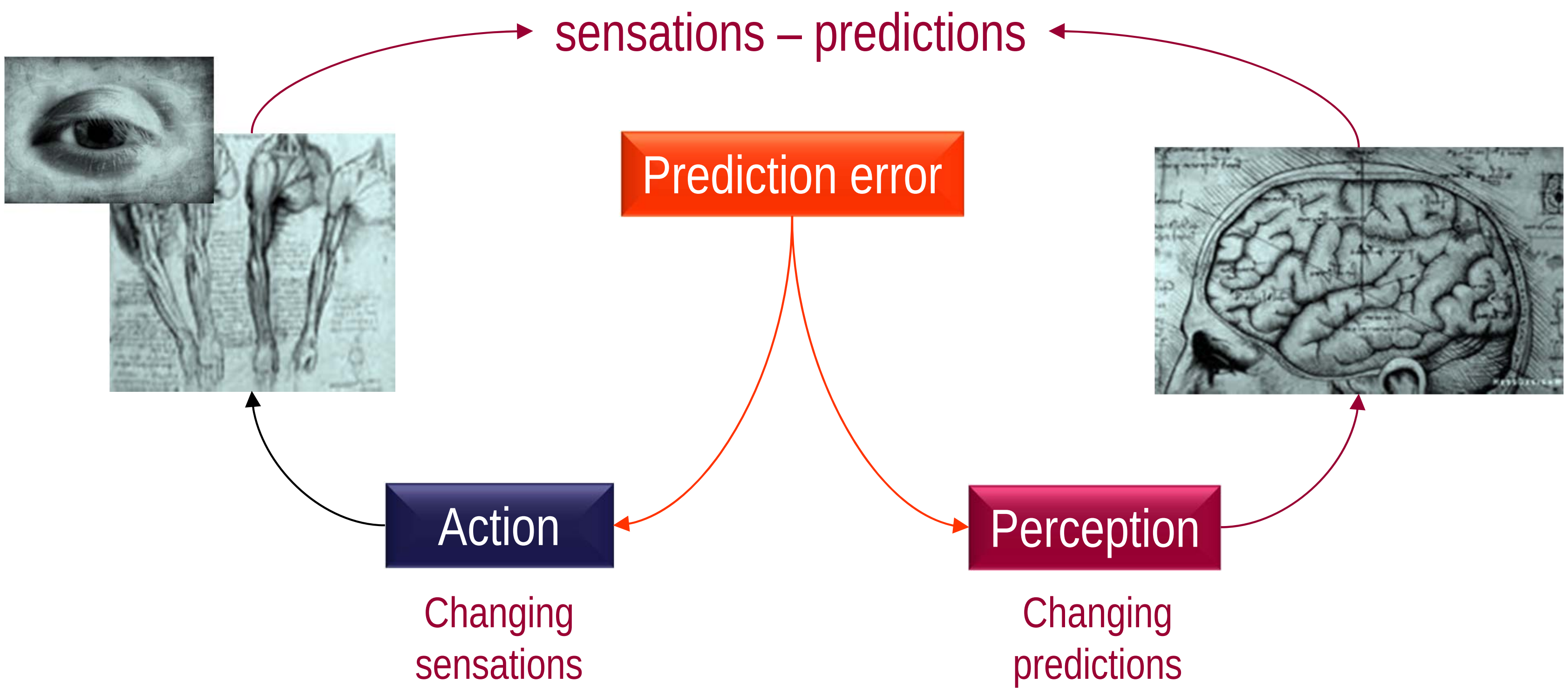
Plato: The Republic (514a-520a)

Bayesian filtering and predictive coding

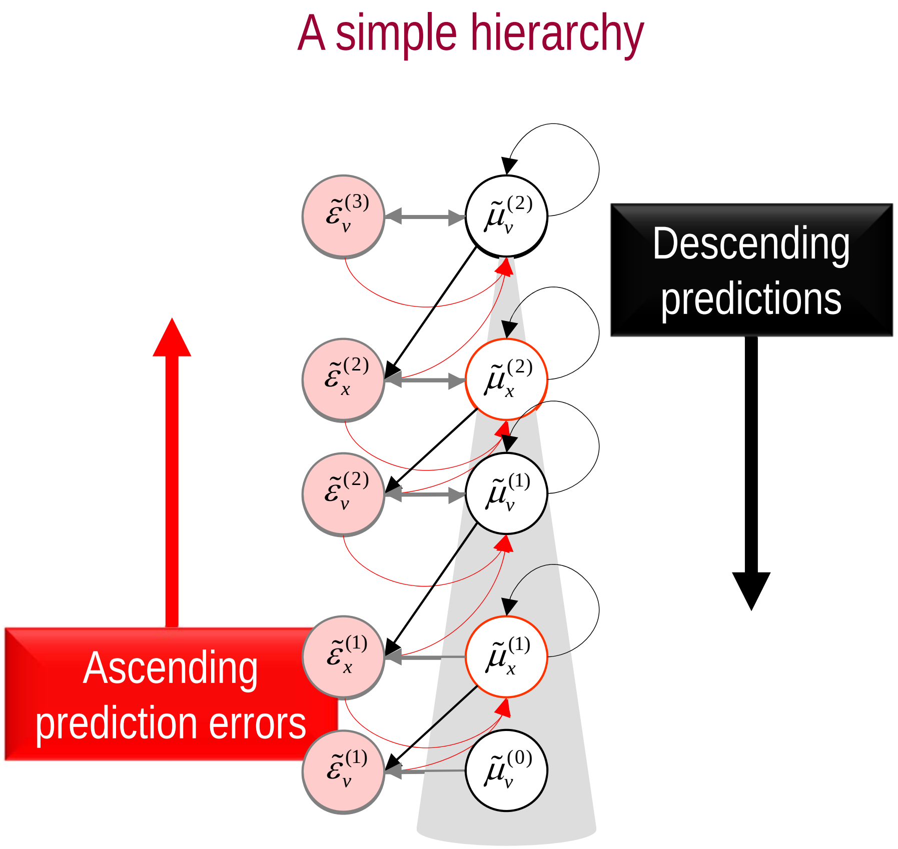
$$
\begin{aligned}
f_{\mu}(\tilde{s}, \tilde{a}, \tilde{\mu}) &= (\Gamma - Q)\nabla \ln p(\tilde{s} \mid m) \\
&= \underbrace{D\tilde{\mu}}_{\text{prediction}} - \underbrace{\Gamma \nabla \tilde{\varepsilon} \cdot \Pi \cdot \tilde{\varepsilon}}_{\text{update}}
\end{aligned}
$$



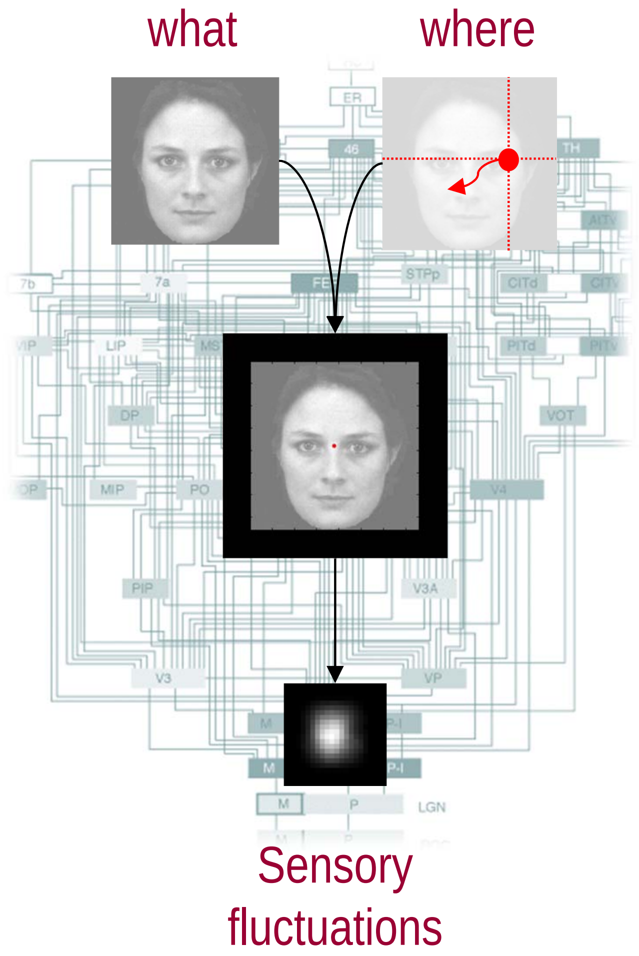
Making our own sensations



Hierarchical generative models



$$\dot{\tilde{\mu}} = D\tilde{\mu} - \Gamma \nabla \tilde{\varepsilon} \cdot \Pi \cdot \tilde{\varepsilon}$$

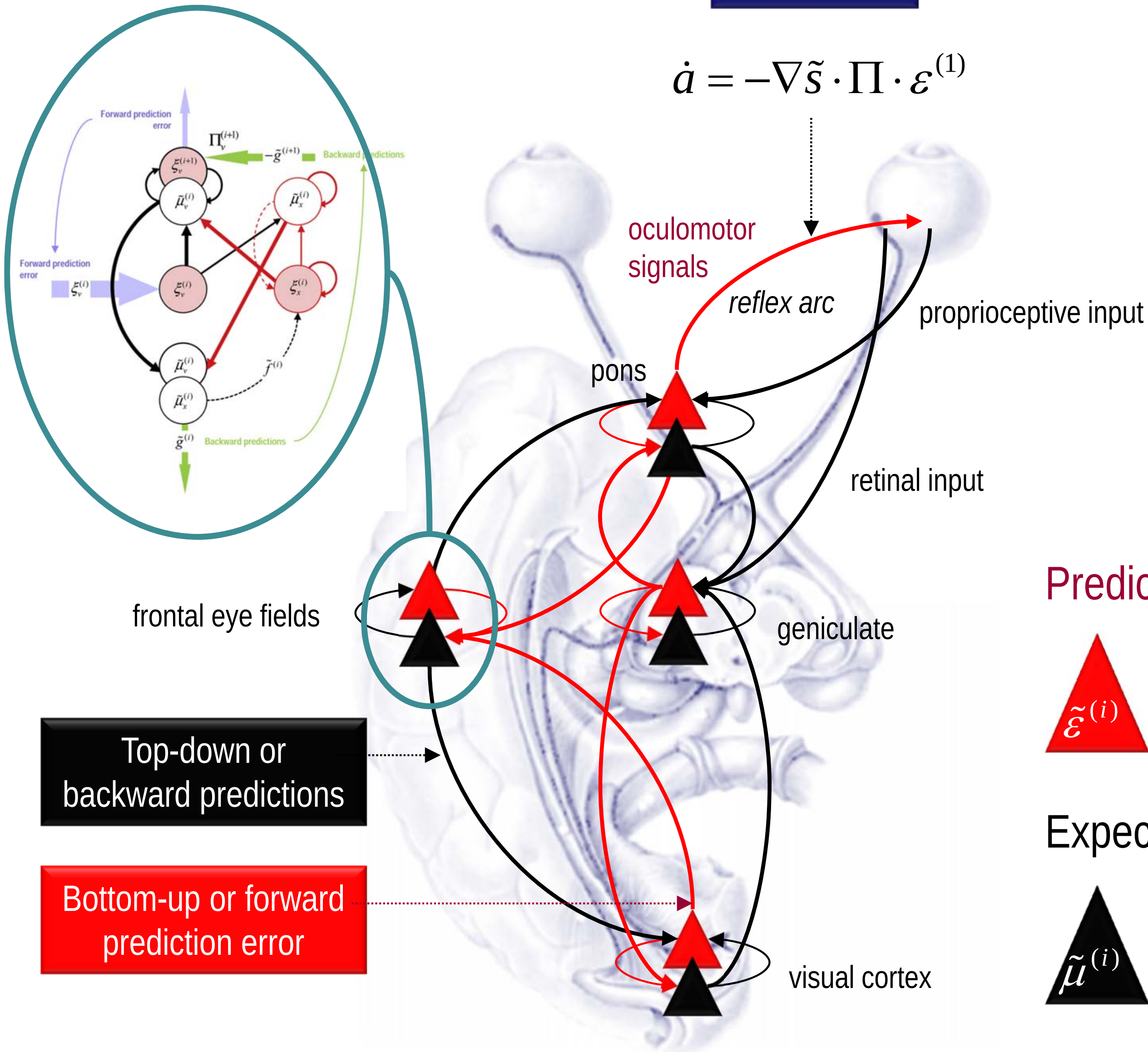


David Mumford

Predictive coding with reflexes

Action

$$\dot{a} = -\nabla\tilde{s}\cdot\Pi\cdot\varepsilon^{(1)}$$



Perception

Prediction error (superficial pyramidal cells)

$\tilde{\varepsilon}^{(i)}$

$$\tilde{\varepsilon}^{(i)} = \tilde{\mu}^{(i-1)} - g^{(i)}(\tilde{\mu}^{(i)})$$

Expectations (deep pyramidal cells)

$\tilde{\mu}^{(i)}$

$$\dot{\tilde{\mu}}^{(i)} = D\tilde{\mu}^{(i)} - \nabla\tilde{\varepsilon}^{(i)}\cdot\Pi^{(i)}\cdot\tilde{\varepsilon}^{(i)}$$

Biological agents minimize their average surprise (entropy)

They minimize surprise by suppressing prediction error

Prediction error can be reduced by changing predictions (perception)

Prediction error can be reduced by changing sensations (action)

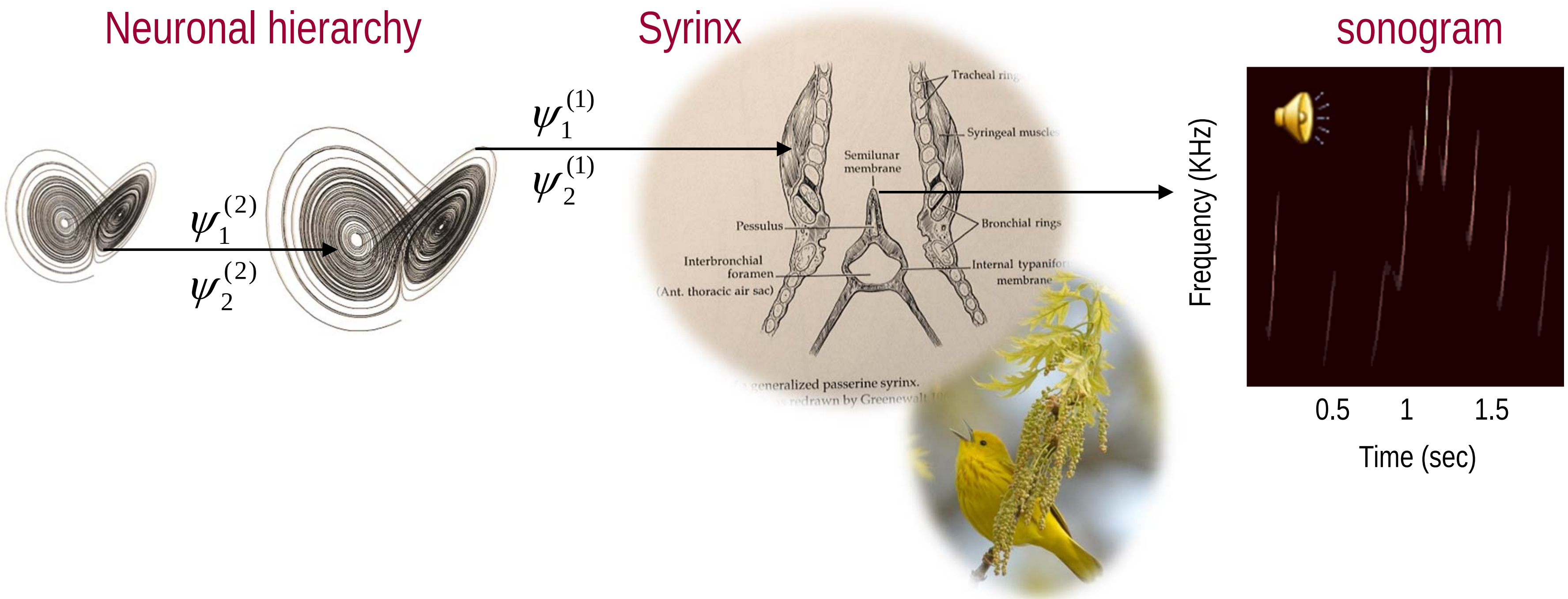
Perception entails recurrent message passing to optimize predictions

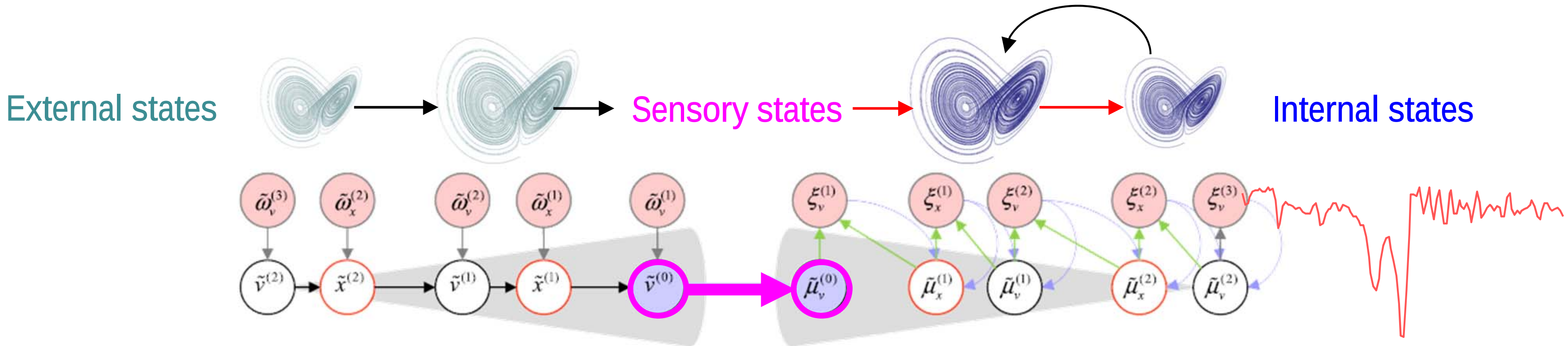
Action makes predictions come true (and minimizes surprise)

Overview

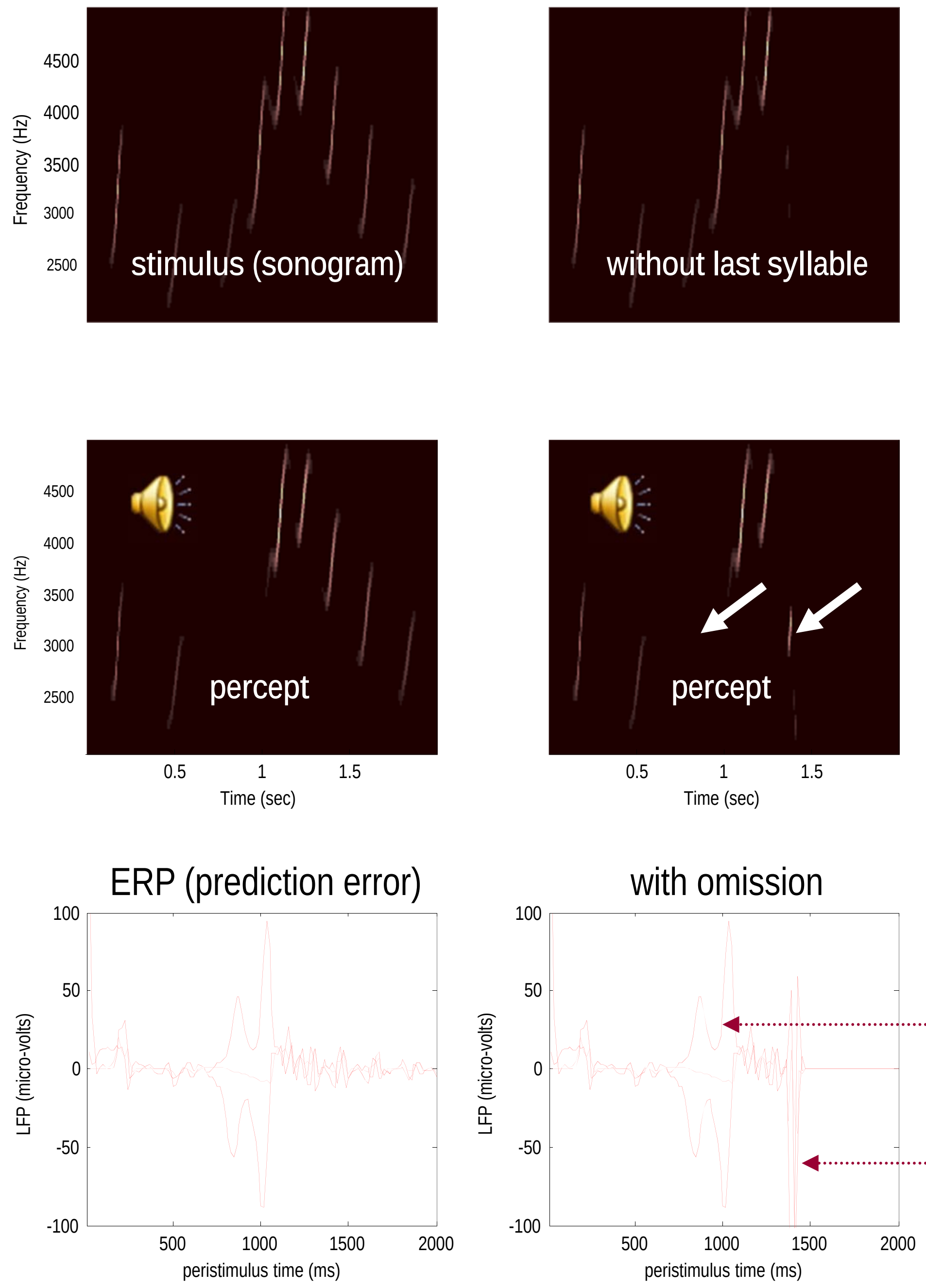
The statistics of life	Markov blankets and ergodic systems simulations of a primordial soup
The anatomy of inference	graphical models and predictive coding canonical microcircuits
Action and perception	perceptual omission responses simulations of action observation

Perceptual inference and sequences of sequences





omission and violation of predictions



Stimulus but no percept

Percept but no stimulus

Overview

The statistics of life

Markov blankets and ergodic systems
simulations of a primordial soup

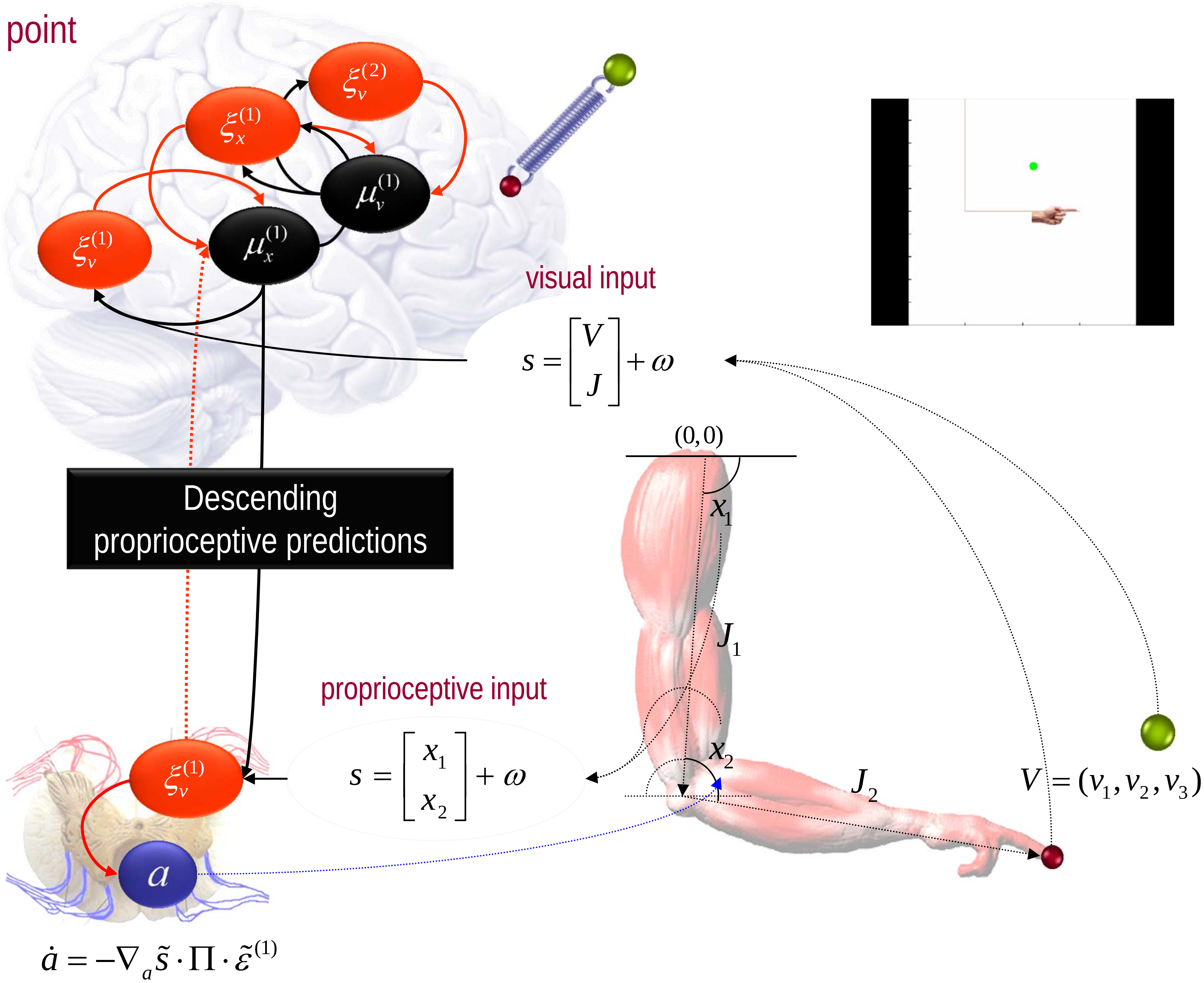
The anatomy of inference

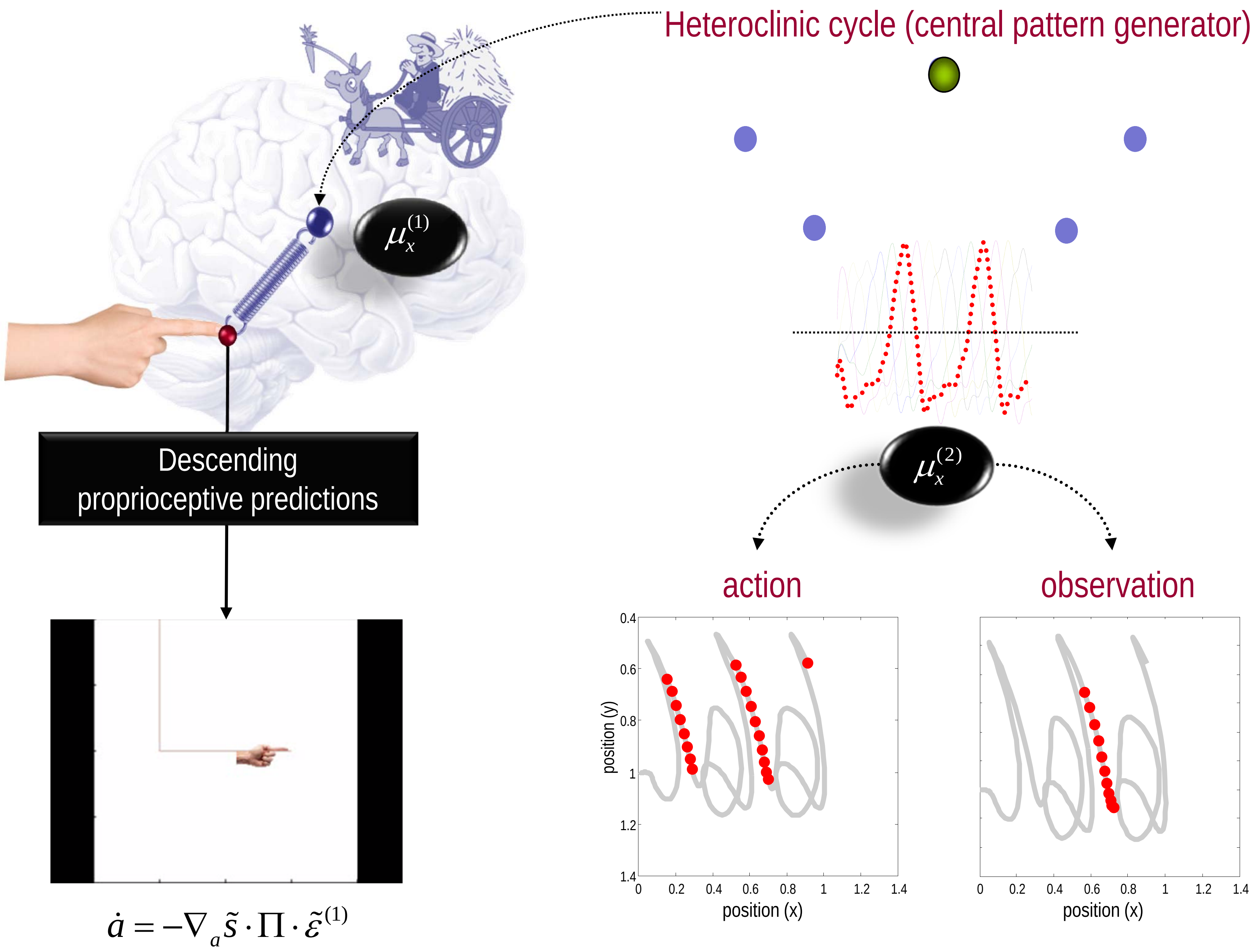
graphical models and predictive coding
canonical microcircuits

Action and perception

perceptual omission responses
simulations of action observation

Action with point attractors





Heteroclinic cycle (central pattern generator)
$\mu_x^{(1)}$
$\mu_x^{(2)}$
Descending proprioceptive predictions
action
observation
position (y)
position (x)
0.4
0.6
0.8
1
1.2
1.4
0
0.2
0.4
0.6
0.8
1
1.2
1.4
$\dot{a} = -\nabla_a \tilde{s} \cdot \Pi \cdot \tilde{\varepsilon}^{(1)}$

Hermann von Helmholtz

"Each movement we make by which we alter the appearance of objects should be thought of as an experiment designed to test whether we have understood correctly the invariant relations of the phenomena before us, that is, their existence in definite spatial relations."

'he Facts of Perception'(1878) in The Selected Writings of Hermann von Helmholtz, Ed. R. Karl, Middletown: Wesleyan University Press, 1971 p. 384

Thank you

And thanks to collaborators:

Rick Adams
Andre Bastos
Sven Bestmann
Harriet Brown
Jean Daunizeau
Mark Edwards
Xiaosi Gu
Lee Harrison
Stefan Kiebel
James Kilner
Jérémie Mattout
Rosalyn Moran
Will Penny
Lisa Quattrocki Knight
Klaas Stephan

And colleagues:

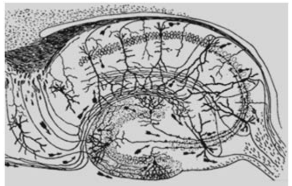
Andy Clark
Peter Dayan
Jörn Diedrichsen
Paul Fletcher
Pascal Fries
Geoffrey Hinton
Allan Hobson
James Hopkins
Jakob Hohwy
Henry Kennedy
Paul Verschure
Florentin Wörgötter

And many others

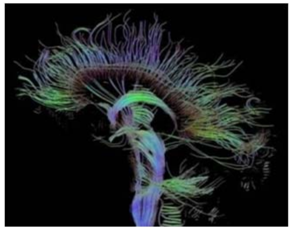
Time-scale

$10^{-3}\,s$

$10^{0}\,s$



$10^{3}\,s$



$10^{6}\,s$

$10^{15}\,s$

Free-energy minimisation leading to…

Perception and Action: The optimisation of neuronal and neuromuscular activity to suppress prediction errors (or free-energy) based on generative models of sensory data.

Learning and attention: The optimisation of synaptic gain and efficacy over seconds to hours, to encode the precisions of prediction errors and causal structure in the sensorium. This entails suppression of free-energy over time.

Neurodevelopment: Model optimisation through activity-dependent pruning and maintenance of neuronal connections that are specified epigenetically

Evolution: Optimisation of the average free-energy (free-fitness) over time and individuals of a given class (e.g., conspecifics) by selective pressure on the epigenetic specification of their generative models.

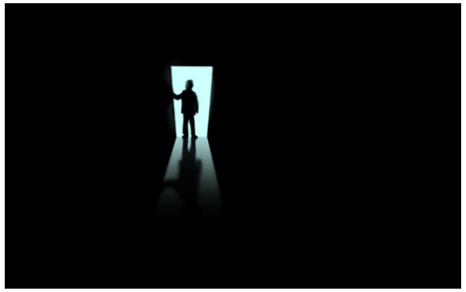
frontiers in
PSYCHOLOGY

HYPOTHESIS AND THEORY ARTICLE
published: 08 May 2012
doi: 10.3389/fpsyg.2012.00130

Free-energy minimization and the dark-room problem

Karl Friston*[1], *Christopher Thornton*[2] and *Andy Clark*[3]

[1] The Wellcome Trust Centre for Neuroimaging, University College London, London, UK
[2] Informatics, University of Sussex, Brighton, UK
[3] School of Philosophy, Psychology, and Language Sciences, University of Edinburgh, Edinburgh, UK



Searching to test hypotheses – life as an efficient experiment

$$
\begin{aligned}
H(S,\Psi) &= H(S\mid m) + H(\Psi\mid S) \\
&= \underbrace{E_t[-\ln p(\tilde{s}(t)\mid m)]}_{\text{Free energy principle}} + \underbrace{E_t[H(\Psi\mid S=\tilde{s}(t))]}_{\text{minimise uncertainty}}
\end{aligned}
$$

$$
\tilde{\eta}(t) = \arg\min_{\tilde{\eta}}\{H[q(\tilde{\psi}\mid\tilde{\mu},\tilde{\eta})]\}
$$