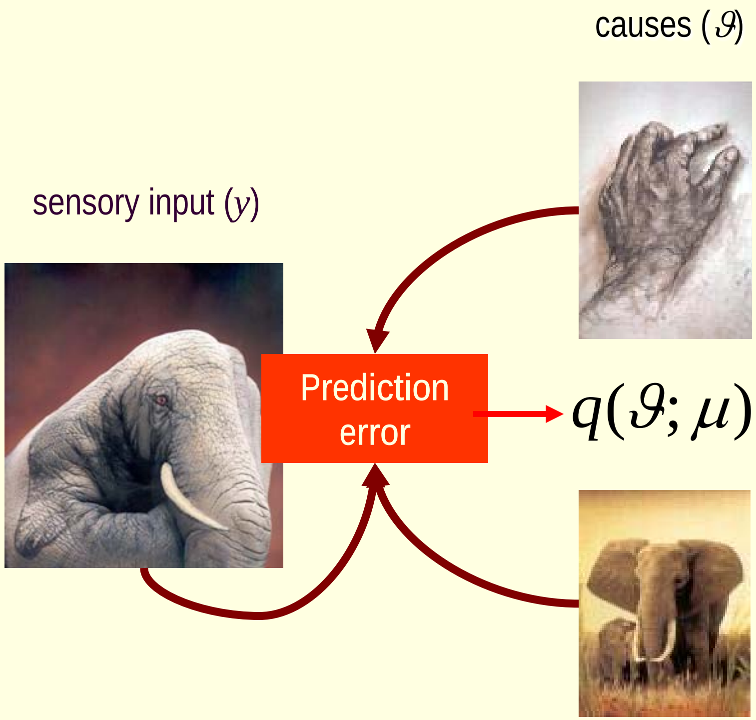
Free energy and value

Collège de France 2008

Abstract

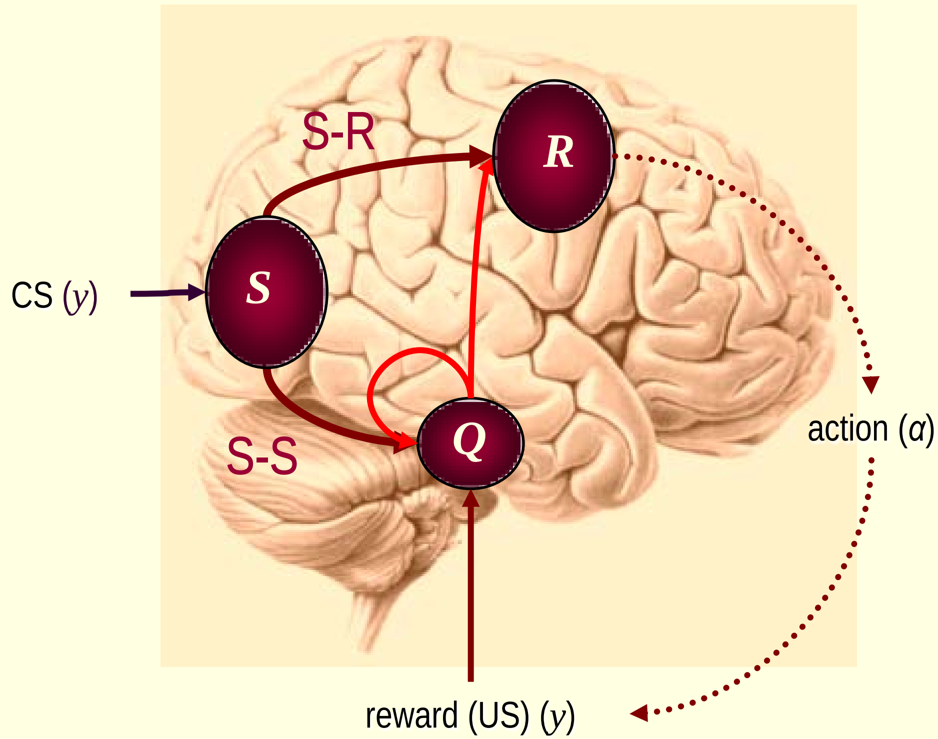
This talk summarizes recent attempts to integrate action and perception within a single optimization framework. We start with a statistical formulation of Helmholtz's ideas about neural energy to furnish a model of perceptual inference and learning that can explain a remarkable range of neurobiological facts. Here, we consider these perceptual processes as just one aspect of systems that conform to a free-energy principle. The free-energy considered here represents a bound on the surprise inherent in exchange with the environment, under expectations encoded by state or configuration of an agent. An agent can minimize free-energy by changing its configuration to change the way it samples the environment, or to change its expectations. These changes correspond to action and perception respectively and lead to active exchange with the environment that is characteristic of biological systems. This treatment implies that the an agents state and structure encode an implicit and probabilistic model of the environment and that its actions suppress surprising exchanges with it. Furthermore, it suggests that free-energy, surprise and value are all the same thing. We will look at models entailed by the brain and how minimization of free-energy can explain its dynamics and structure.

Perceptual learning



$$\min_{\mu} D(q(\vartheta;\mu)\,||\,p(\vartheta\,|\,y))$$

Value learning

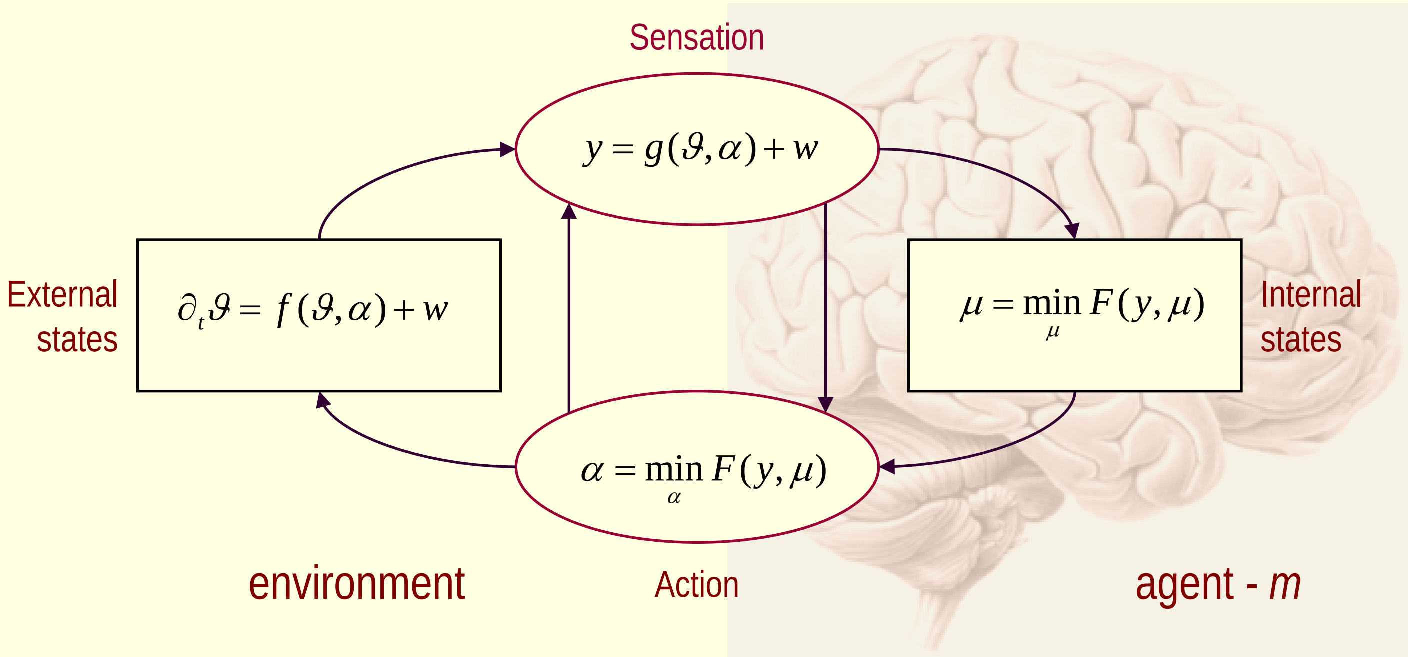


$$\max_{\alpha} Q(y\,|\,\vartheta,\alpha)$$

Overview

Surprise, value and ensemble densities
The free energy principle
Perception with hierarchical models
Seeing a smile

Exchange with the environment



Separated by a Markov blanket

Selection and exchanges with the environment

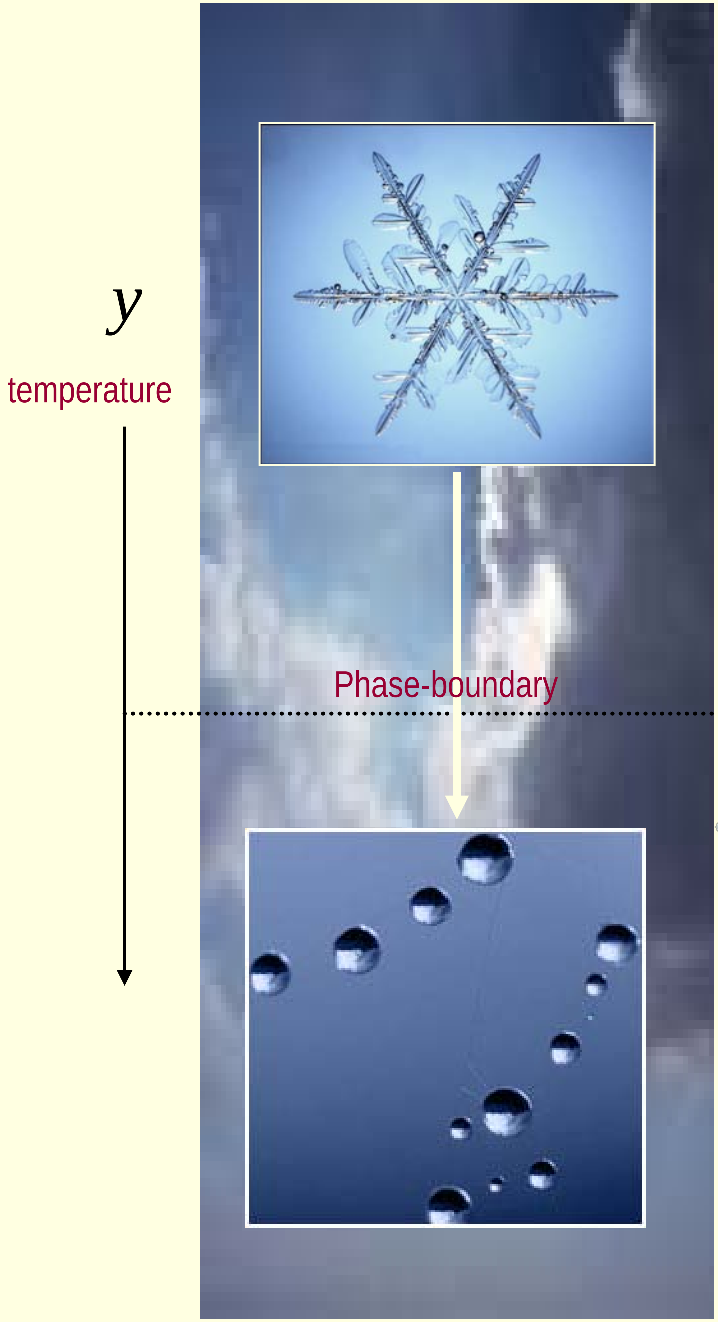
Darwinian evolution of virtual block creatures. A population of several hundred creatures is created within a supercomputer, and each creature is tested for their ability to perform a given task, such the ability to swim in a simulated water environment. The successful survive, and their virtual genes are copied, combined, and mutated to make offspring. The new creatures are again tested, and some may be improvements on their parents. As this cycle of variation and selection continues, creatures with more and more successful behaviours can emerge.

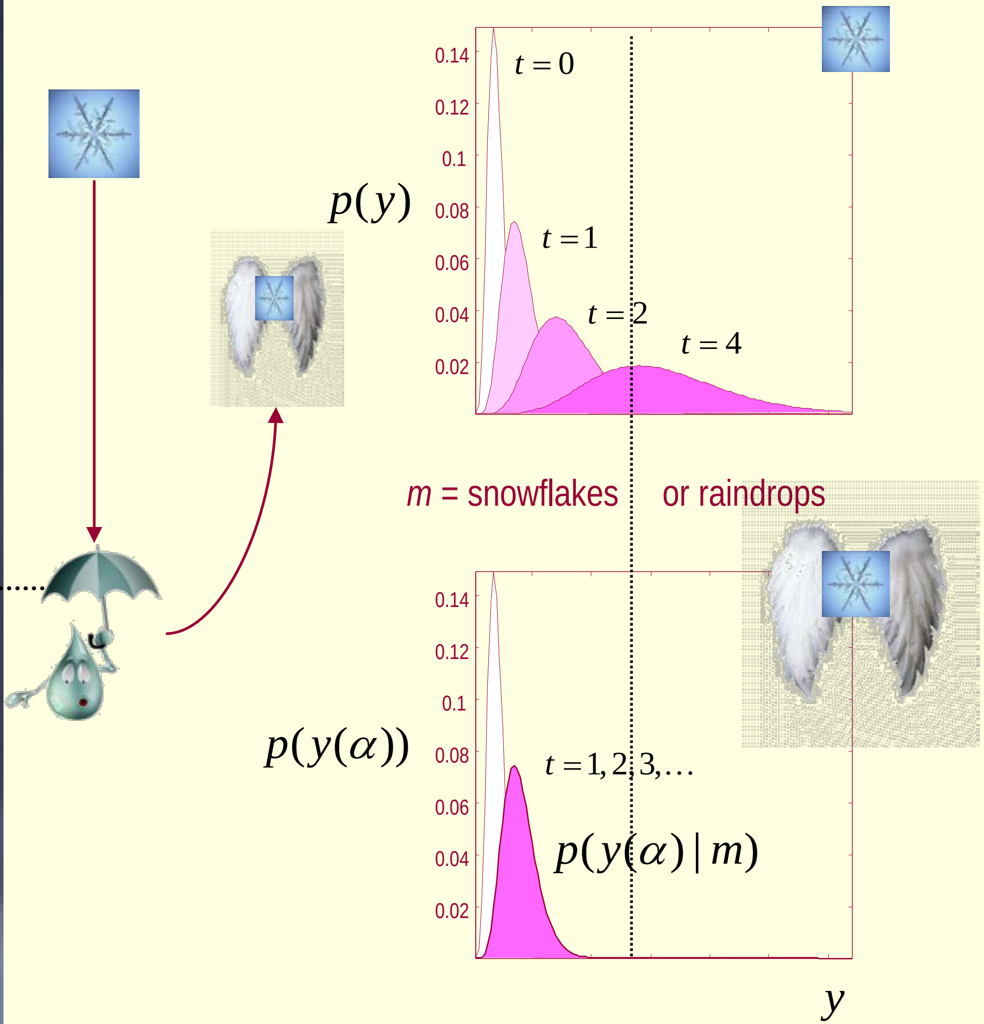
What we are, m is defined by our exchange with the environment y;
i.e., by the equilibrium density

$$p(y\,|\,m)$$

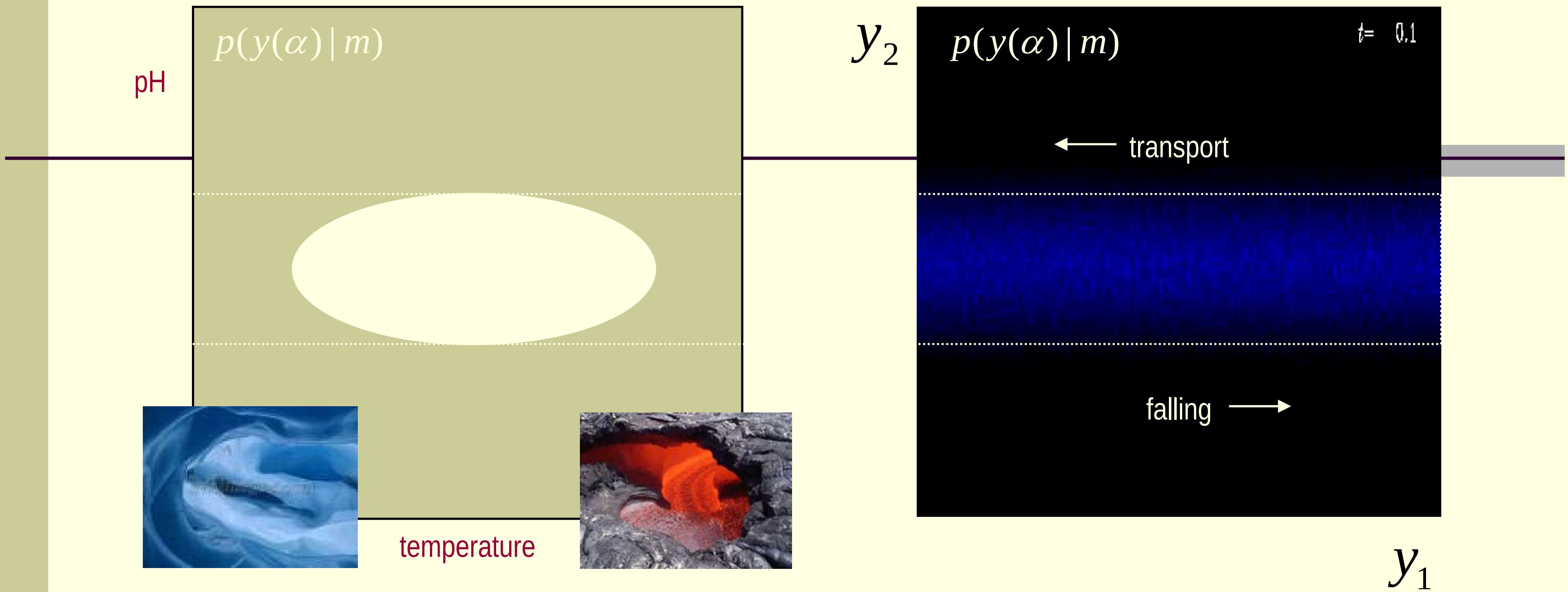
Beyond neo-classical selection: active agents …

Open systems and active agents





Active agents maximise value (or minimize surprise): $\ln p(y(\alpha) \mid m) = Q$



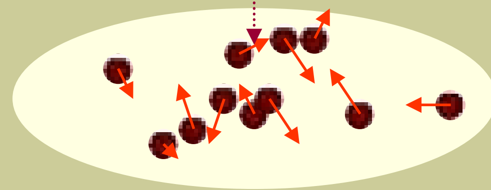
Self-organization and equilibria

Particle density contours showing Kelvin-Helmholtz instability, forming beautiful breaking waves. In the self-sustained state of Kelvin-Helmholtz turbulence the particles are transported away from the mid-plane at the same rate as they fall, but the particle density is nevertheless very clumpy because of a clumping instability that is caused by the dependence of the particle velocity on the local solids-to-gas ratio (Johansen, Henning, & Klahr 2006)

$y_2 = pH$

$p(y(\alpha)\,|\,m)$

$$\frac{\partial p}{\partial t} = -\nabla \cdot [\, f(\alpha)\, p - D\nabla p\,]$$



Ensemble dynamics

y_1

At equilibrium, action maximises value or minimises surprise

$$\frac{\partial p}{\partial t} = 0 \Leftarrow f(\alpha)\, p = D\nabla p \Rightarrow f(\alpha) = D\frac{\partial Q}{\partial y}$$

$$Q = \ln p(y(\alpha))$$

Beyond self-organization

Ensemble dynamics and swarming

One small problem…

Agents cannot access Q. However they can evaluate a free-energy bound on surprise

$$-\ln p(y(\alpha)\,|\,m) = -Q \leq F$$

i.e., under equilibrium, active agents optimise their free-energy

Overview

Surprise, value and ensemble densities
The free energy principle
Perception with hierarchical models
Seeing a smile

The free-energy principle

$$F = -\langle \ln p(y(\alpha), \vartheta \mid m) \rangle_q + \langle \ln q(\vartheta) \rangle_q \geq -Q$$

Action to minimise a bound on surprise

$$F = -\langle \ln p(y(\alpha) \mid \vartheta, m) \rangle_q + D(q \,\|\, p(\vartheta))$$

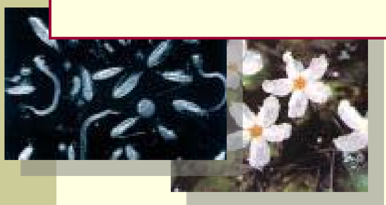
$$\begin{aligned} \alpha &= \min_{\alpha} F \\ &= \max_{\alpha} \langle \ln p(y(\alpha) \mid \vartheta, m) \rangle_q \end{aligned}$$

Perception to optimise the bound

$$F = -\ln p(y \mid m) + D(q(\vartheta;\mu) \,\|\, p(\vartheta \mid y))$$

$$\mu = \min_{\mu} F \Rightarrow$$

$$q(\vartheta;\mu) \to p(\vartheta \mid y)$$



$$\vartheta = \{u, \theta, \gamma\}$$

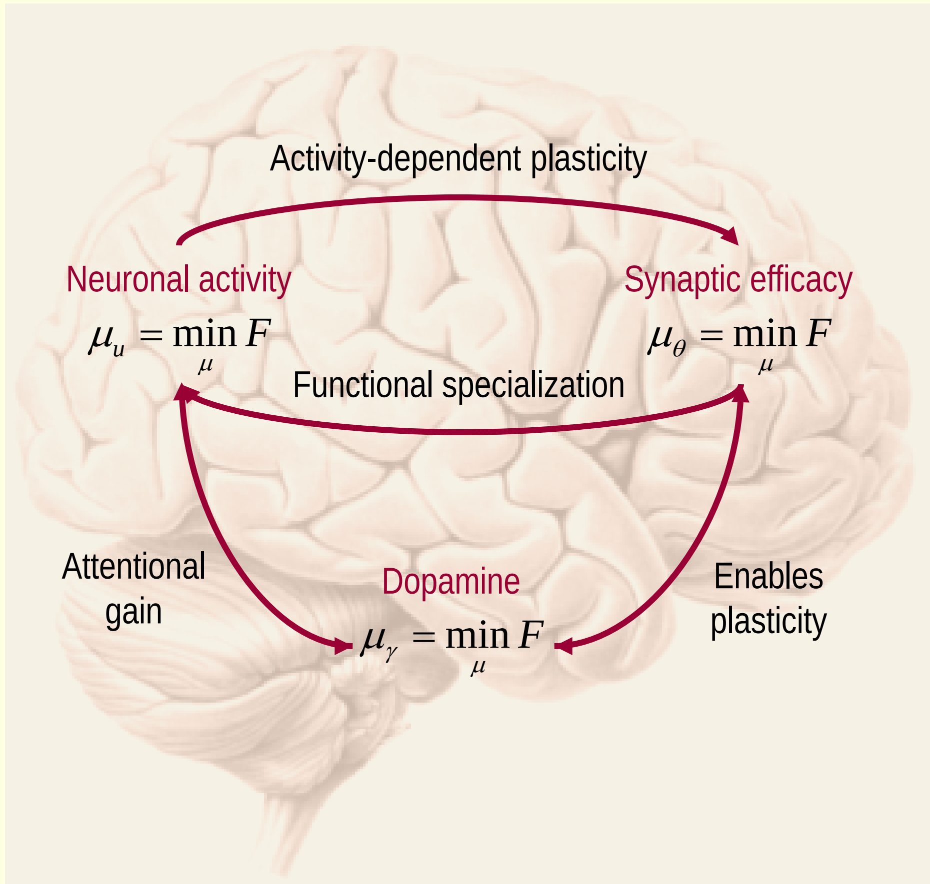
The conditional density and separation of scales

$$q(\vartheta;\mu) = q(u;\mu_u)\,q(\theta;\mu_\theta)\,q(\gamma;\mu_\gamma)$$

Perceptual inference	Perceptual learning	Perceptual uncertainty
$\mu_u = \min_{\mu} F$	$\mu_\theta = \min_{\mu} F$	$\mu_\gamma = \min_{\mu} F$

Mean-field interactions and the brain

$$q(\vartheta;\mu) = q(u;\mu_u)q(\theta;\mu_\theta)q(\gamma;\mu_\gamma)$$



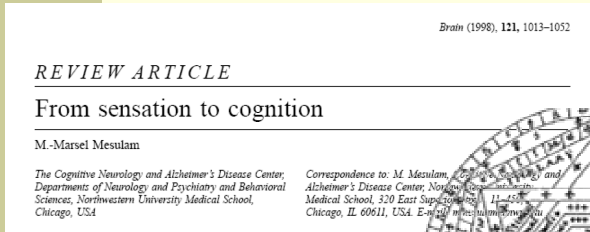
Overview

Surprise, value and ensemble densities
The free energy principle
Perception with hierarchical models
Seeing a smile

Active agents optimise their free-energy to maintain their equilibrium density in a changing environment. Free-energy entails a model of the environment that is parameterized by the agents state and structure. What do these models look like?

… from principles to models

Neuronal hierarchies and hierarchical models


Brain (1998), 121, 1013–1052

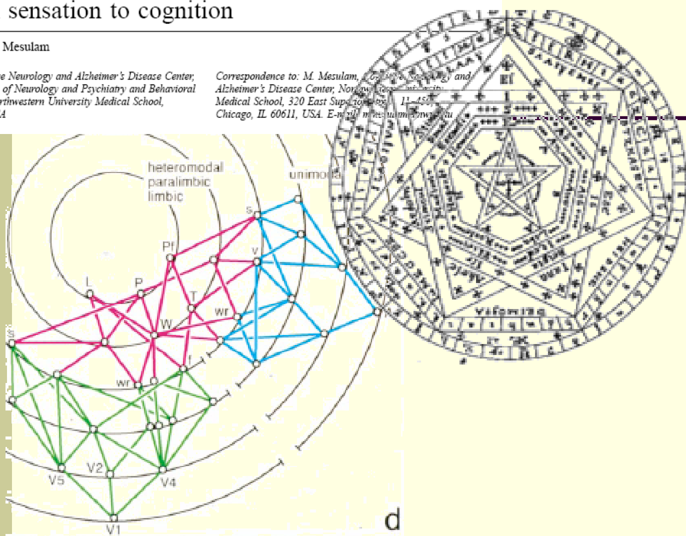
REVIEW ARTICLE

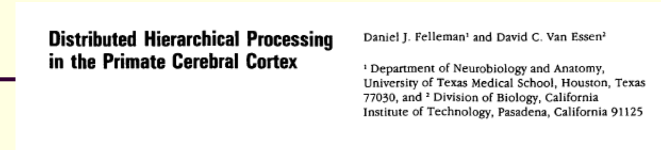
From sensation to cognition

M.-Marsel Mesulam

The Cognitive Neurology and Alzheimer's Disease Center, Departments of Neurology and Psychiatry and Behavioral Sciences, Northwestern University Medical School, Chicago, USA

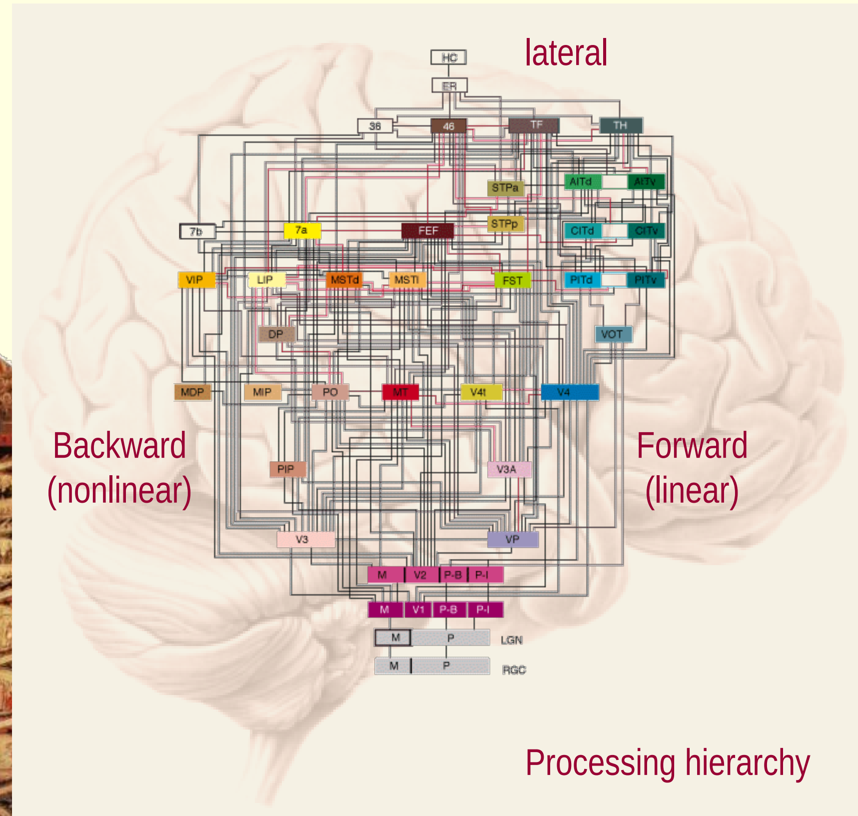
Correspondence to: M. Mesulam, Cognitive Neurology and Alzheimer's Disease Center, Northwestern University Medical School, 320 East Superior ... Chicago, IL 60611, USA. E-mail ...

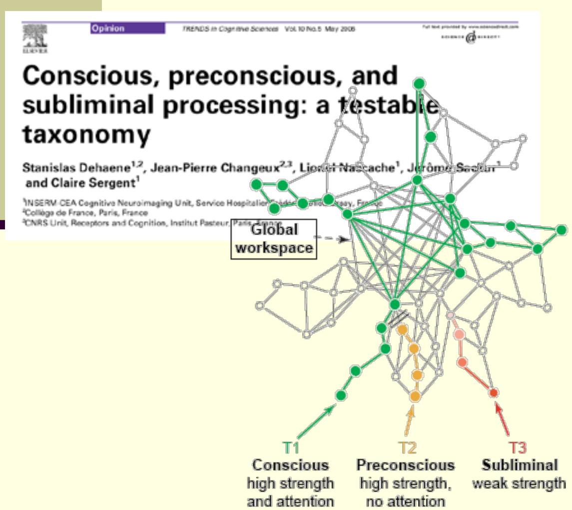



Distributed Hierarchical Processing in the Primate Cerebral Cortex

Daniel J. Felleman[1] and David C. Van Essen[2]

[1] Department of Neurobiology and Anatomy, University of Texas Medical School, Houston, Texas 77030, and [2] Division of Biology, California Institute of Technology, Pasadena, California 91125




Opinion

TRENDS in Cognitive Sciences Vol.10 No.5 May 2006

Conscious, preconscious, and subliminal processing: a testable taxonomy

Stanislas Dehaene[1,2], Jean-Pierre Changeux[2,3], Lionel Naccache[1], Jérôme Sackur[1] and Claire Sergent[1]

[1] INSERM-CEA Cognitive Neuroimaging Unit, Service Hospitalier ..., France
[2] Collège de France, Paris, France
[3] CNRS Unit, Receptors and Cognition, Institut Pasteur, Paris, France

Hierarchical models and their inversion

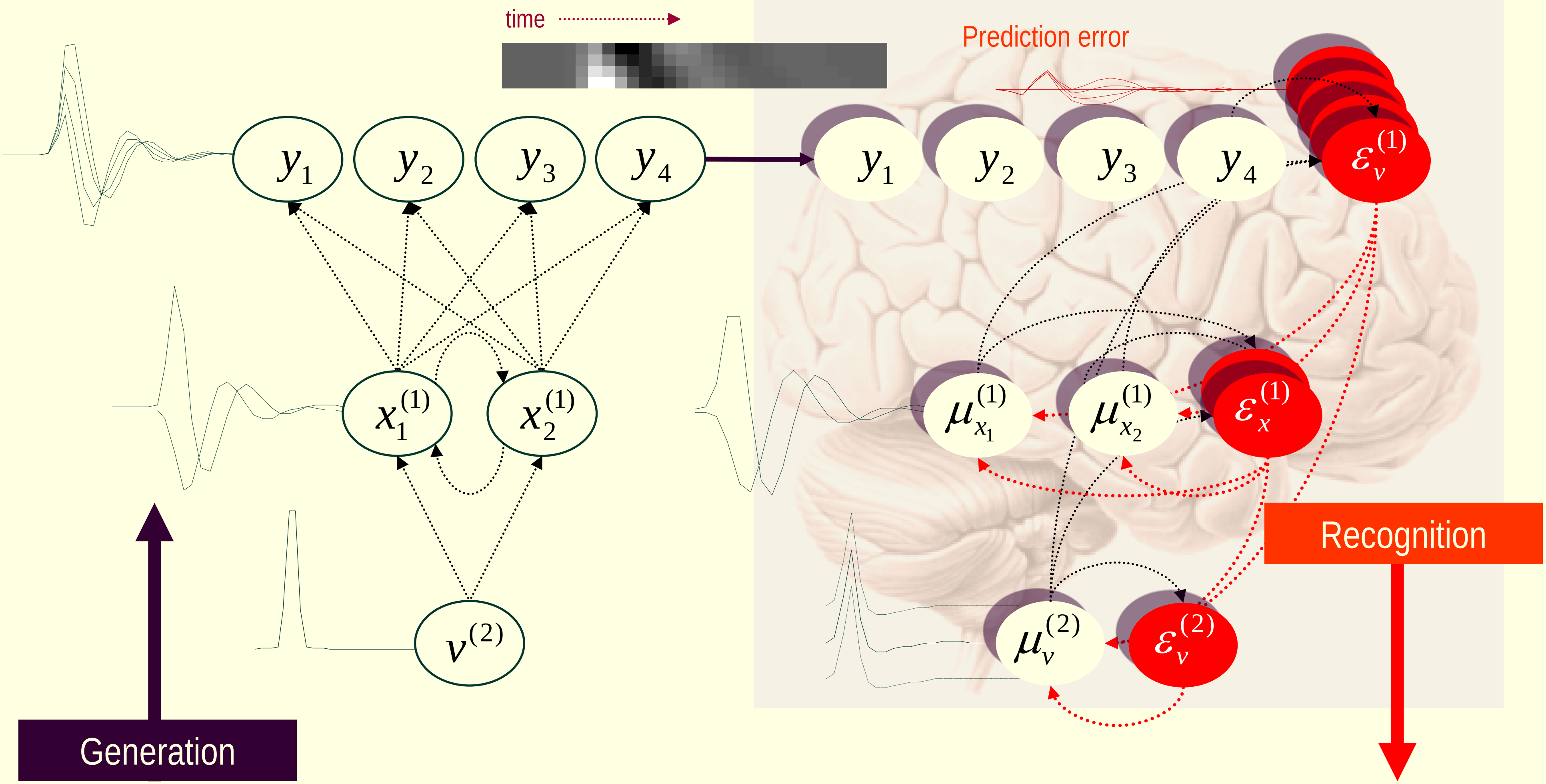
$$v^{(i)} = g(x^{(i)}, v^{(i+1)}) + \varepsilon_v^{(i)}$$
$$\partial_t x^{(i)} = f(x^{(i)}, v^{(i+1)}) + \varepsilon_x^{(i)}$$

Top-down messages

$$\varepsilon_v^{(i)} = \mu_v^{(i)} - g(\mu_x^{(i)}, \mu_v^{(i+1)})$$
$$\varepsilon_x^{(i)} = \partial_t \mu_x^{(i)} - f(\mu_x^{(i)}, \mu_v^{(i+1)})$$

Bottom-up messages

$$\partial_t \mu_v^{(i)} = -\partial_{v\varepsilon^{(i-1)}} F \varepsilon^{(i-1)} - \partial_{v\varepsilon^{(i)}} F \varepsilon^{(i)}$$
$$\partial_t \mu_x^{(i)} = -\partial_{x\varepsilon^{(i)}} F \varepsilon^{(i)}$$



Empirical Bayes and hierarchical models

D-Step
Perceptual inference

Bottom-up

Lateral

$$\partial_t \mu_v^{(i)} = -\partial_v \varepsilon_u^{(i-1)T} \mu_\gamma^{(i-1)} \varepsilon_u^{(i-1)} - \mu_\gamma^{(i)} \varepsilon_u^{(i)}$$

$$\varepsilon_v^{(i)} = \mu_v^{(i)} - g(\mu_x^{(i)}, \mu_v^{(i+1)})$$

Top-down

The balance between top-down empirical priors and bottom-up prediction error is controlled by the efficacy of lateral interactions encoding uncertainty or precision; *classical neuromodulation*

E-Step
Perceptual learning

$$\partial_t \mu_\theta^{(i)} = -\left\langle \partial_\theta \varepsilon_u^{(i)T} \mu_\gamma^{(i-1)} \varepsilon_u^{(i)} \right\rangle$$

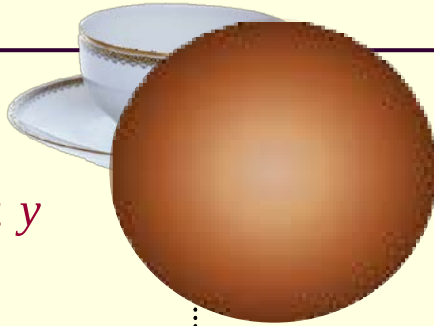
Enabling of associative plasticity; *a permissive role in modulating synaptic plasticity*

M-Step
Perceptual uncertainty

$$\partial_t \mu_\gamma^{(i)} = \mu_\gamma^{(i)-1} - \varepsilon_u^{(i)T} \varepsilon_u^{(i)}$$

Friston K Kilner J Harrison L A free energy principle for the brain. J Physiol Paris. 2006

Perception and message passing



sensory input y

AMPA

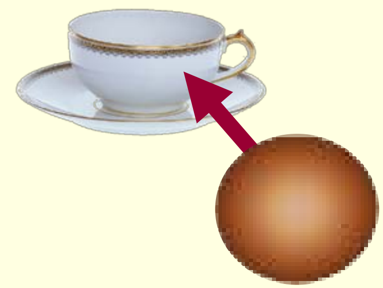
$$\partial_t \mu_v^{(i)} = -\partial_{v\varepsilon^{(i-1)}} F \varepsilon^{(i-1)} -$$

Forward connections convey *feedback*

$$\varepsilon = y - g(\mu)$$

Prediction error

causes μ



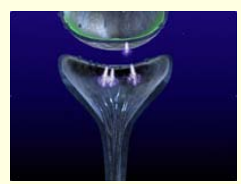
$$g(\mu)$$

Prediction

Backward connections convey the *forward* model

$$\varepsilon_v^{(i)} = \mu_v^{(i)} - g(\mu_x^{(i)}, \mu_v^{(i+1)})$$

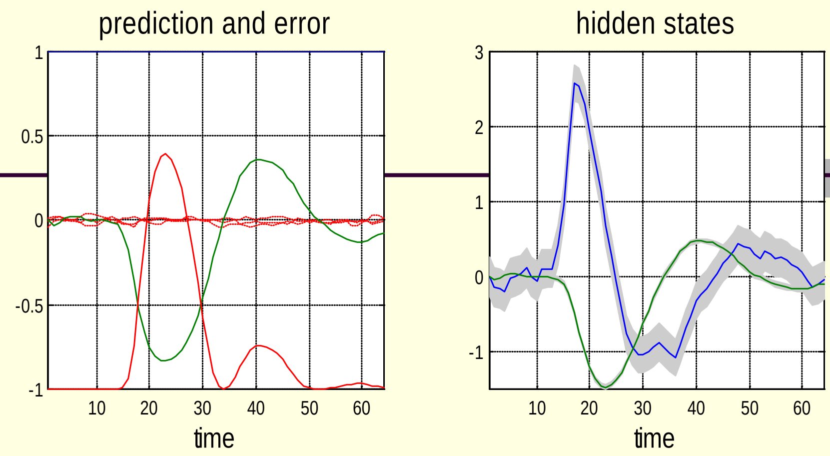
NMDA

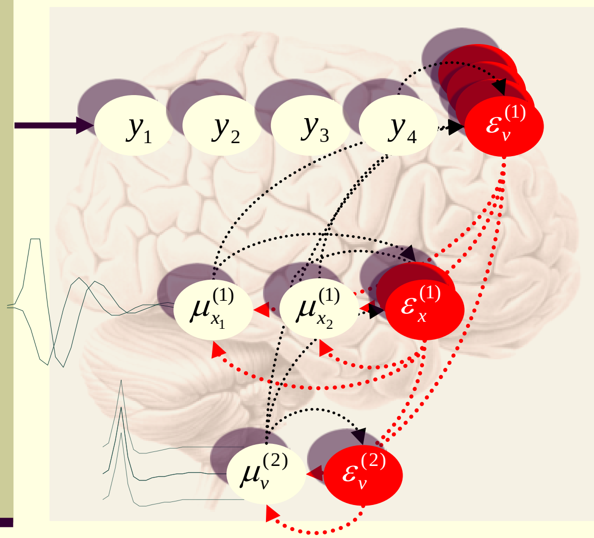


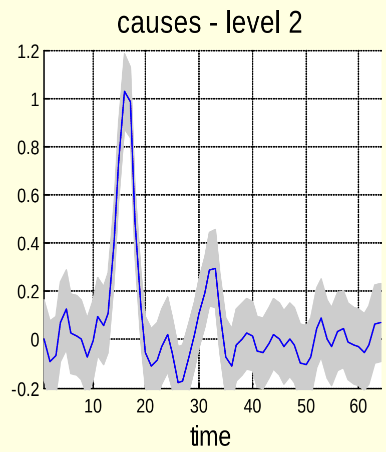
Overview

Surprise, value and ensemble densities
The free energy principle
Perception with hierarchical models
Seeing a smile

Bayesian inversion of visual transients



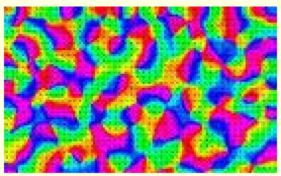




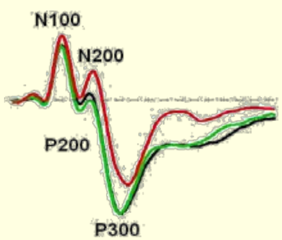
Predictions about the brain

This model of brain function explains a wide range of anatomical and physiological facts; for example, the hierarchical deployment of cortical areas, recurrent architectures using forward and backward connections and functional asymmetries in these connections (Angelucci et al, 2002).

In terms of synaptic physiology, it predicts associative plasticity and, for dynamic models, spike–timing-dependent plasticity.



In terms of electrophysiology it accounts for classical and extra-classical receptive field effects and long-latency or endogenous components of evoked cortical responses (Rao and Ballard, 1998).



It predicts the attenuation of responses encoding prediction error, with perceptual learning, and explains many phenomena like repetition suppression, mismatch negativity and the P300 in electroencephalography.

In psychophysical terms, it accounts for the behavioural correlates of these physiological phenomena, *e.g.*, priming, and global precedence.

Summary

- A free energy principle can account for several aspects of action and perception
- The architecture of cortical systems speak to hierarchical generative models
- Estimation of hierarchical dynamic models corresponds to a generalised deconvolution of inputs to disclose their causes
- This deconvolution can be implemented in a neuronally plausible fashion by constructing a dynamic system that self-organises when exposed to inputs to suppress its free energy