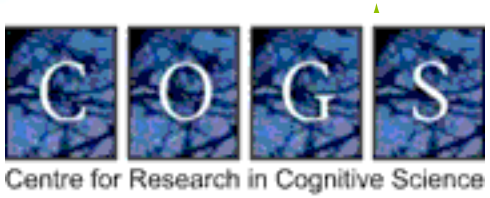
Making Predictive Coding More Predictive, More Enactive

Ron Chrisley

Sackler Centre for Consciousness Science
Centre for Research in Cognitive Science
PAICS Lab
School of Engineering and Informatics
University of Sussex

Presented at the 16th annual meeting of the Association for the Scientific Study of Consciousness
July 3rd, 2012, Brighton, UK

Sackler Centre for Consciousness Science



Overview

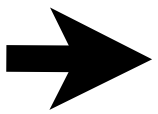
- The appeal of Predictive Coding (PC) models
  - *Thank you, prior speakers*
- Challenges of applying such models to consciousness
- Offer one generalisation of PC models, the Expectational Model (EM), as a way of:
  - a) identifying these challenges
  - b) suggesting some ways to overcome them
  - c) identifying which features of PC models facilitate this and which do not

From predictive coding models to an expectation-based model

Features retained, transformed, demoted

Typical PC model

- Predictive
- Error-minimising
- Inferential (Helmholtz)
- Hierarchical (priors)
- Action-involving
- Probabilistic
- Optimal (Bayesian)
- Subtractive ("suppressive")



Expectational Model

- **Predictive**
- **Error-minimising**
- Holistic (Merleau-Ponty)
- Hierarchical (features)
- Enactive
- Probabilistic
- Optimal (Bayesian)
- Subtractive ("suppressive")

Predictive Coding Models of Consciousness

E.g., models of:

- Binocular rivalry (Hohwy, Roepstorff & Friston, 2008)
- Phenomenal presence (Seth, Suzuki & Critchley, 2011)

Also, following Clark's very thorough review and analysis of predictive coding*:

- Delusions and hallucinations in schizophrenia (Fletcher and Frith (2009), Corlett, Frith, and Fletcher (2009))
- Cross- and multi-modal context effects on early sensory processing (Murray et al (2006), Muckli et al (2005) and Muckli (2010), Kriegstein and Giraud (2006), Langner et al (2011))

**Clark, A. "Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science" Behavioral and Brain Sciences (2012, in press)*

Clark on predictive coding

Clark, A. "Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science" *Behavioral and Brain Sciences* (2012, in press)

- Very thorough review and analysis of predictive coding
- Problems for PC models in general
  - E.g., the darkroom problem (Friston, Thornton and Clark 2012)
- Problems for PC models of consciousness in particular
  - In general what is the relation between sub-personal models and personal-level experience?
    - "To what extent, if any, do these [PC] stories capture or explain facts about what we might think of as personal (or agent-level) cognition - the flow of thoughts, reasons, and ideas that characterize daily conscious thought and reason?"

1: Surprise vs Surprisal

- The problem:
  - "[T]here seems to be a large disconnect between 'surprisal' (the implausibility of some sensory state given a model of the world...) and agent-level surprise. This is evident from the simple fact that the percept that, overall, best minimizes surprisal (hence minimizes prediction errors) 'for' the brain may well be, for me the agent, some highly surprising and unexpected state of affairs – for example, the sight of a large pink rabbit (to borrow a striking image from John Haugeland) dancing in the middle of the room."
- Clark's solution?
  - "Given the right driving signal and a high enough assignment of precision, top-level theories of an initially agent-unexpected kind can still win out so as to explain away that highly-weighted tide of incoming sensory evidence."
  - Although this can explain where the "large pink rabbit" component of personal-level experience comes from, it doesn't explain where the "surprising" component of personal level experience comes from.
- By contrast, EM can explain both, and can thus account both for cases of change perception and change blindness.

2: Indeterminacy of PC models vs determinacy of experience

- "The world, it might be said, does not look as if it is encoded as an intertwined set of probability density distributions! It looks unitary and, on a clear day, unambiguous. But this phenomenology again poses no real challenge. What is on offer, after all, is a story about the brain's way of encoding information about the world. It is not directly a story about how things seem to agents deploying that means of encoding information. There is clearly no inconsistency in thinking that the brain's pervasive use of probabilistic encoding might yield conscious experiences that depict a single, unified and quite unambiguous scene."
- To say that there is no inconsistency is a weak defence: If PC models are to explain consciousness, it needs to be shown how personal-level determinacy of experience *follows from* or is *explained by* sub-personal indeterminacy.

EM: An expectational model of consciousness

EM has two parts: An expectation-based architecture (EBA), and an expectation-based theory of consciousness

Roughly:

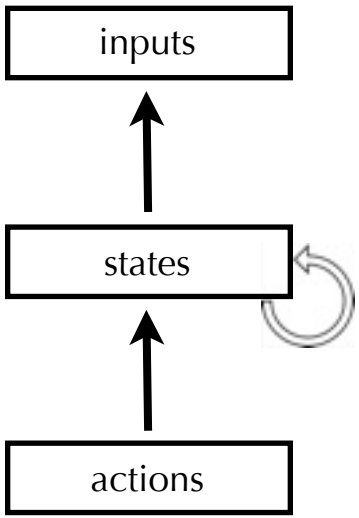
- The *architecture* includes expectations in the form of a forward model: "How will my visual input change if I execute this or that motor command (e.g., eye saccades)?"
- The *theory* posits that the content of visual experience is equal to the content of the expectational state, spatially structured according to the actions ranged over in the forward model

Expectation-based architecture

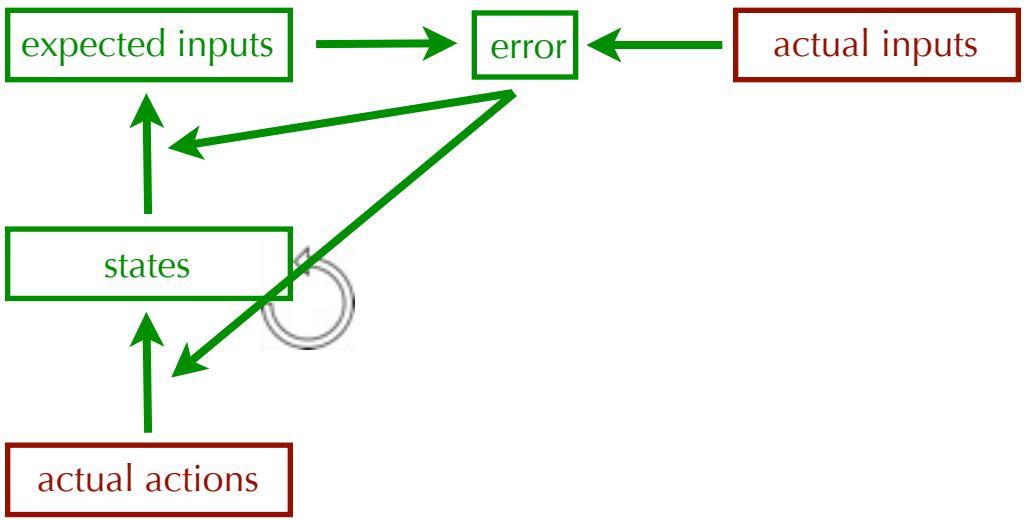
- Only makes sense to talk on expectations in a system that is capable of both perception and action: EBA
  - (Chrisley and Parthemore 2007a; 2007b); Extension of the CNM model (Chrisley 1990)
- At the heart of EBA is a learnable forward model
  - Learnable because it uses powerful supervised learning algorithms without the need for a supervising "teacher"
  - Achieved by re-construing the cognitive "task" to be one for which the world *can* provide the right answers: prediction of future inputs
  - Enactive: prediction of inputs that will be obtained *if a particular action is performed*
  - Learning consists of altering parameters to minimise prediction error
- No separate modules for scene recognition and action-selection
  - Rather, all three modes of cognitive activity are the result of selective modulation and redeployment of the same forward model

EBA as the re-application of a forward model

- The forward model (can be implemented as a simple recurrent network)



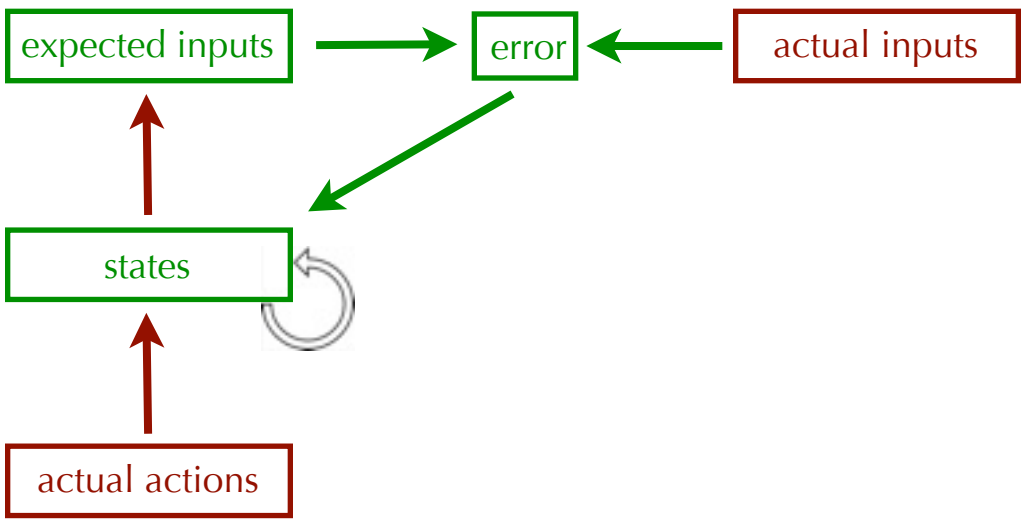
Learning the forward model



- Key: Frozen/fixed by the world; Variable (error minimisation)

Scene recognition as error minimisation

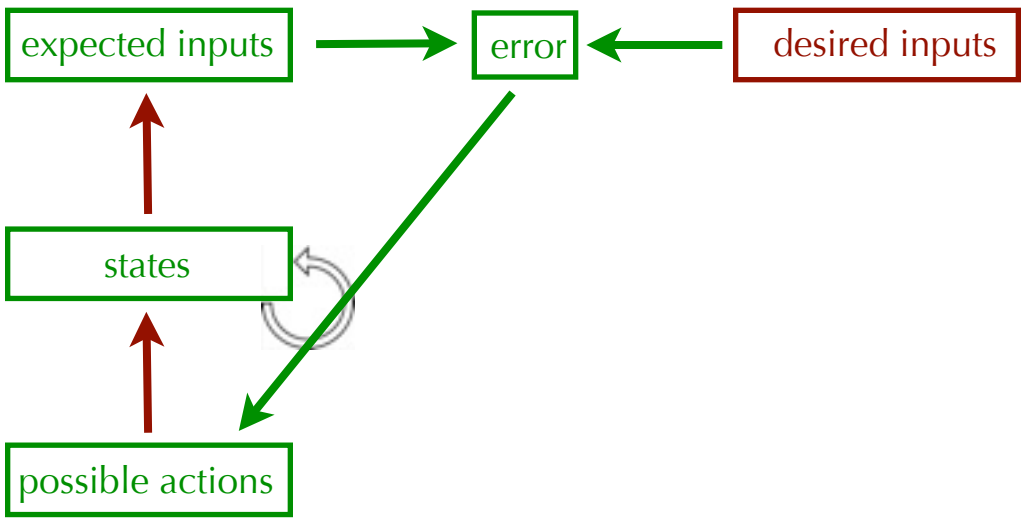
- Fixing the learned forward model to estimate state (e.g., recognise gists, scenes, objects)



- Key: Frozen/fixed by the world; Variable (error minimisation)

Planning as error-minimisation

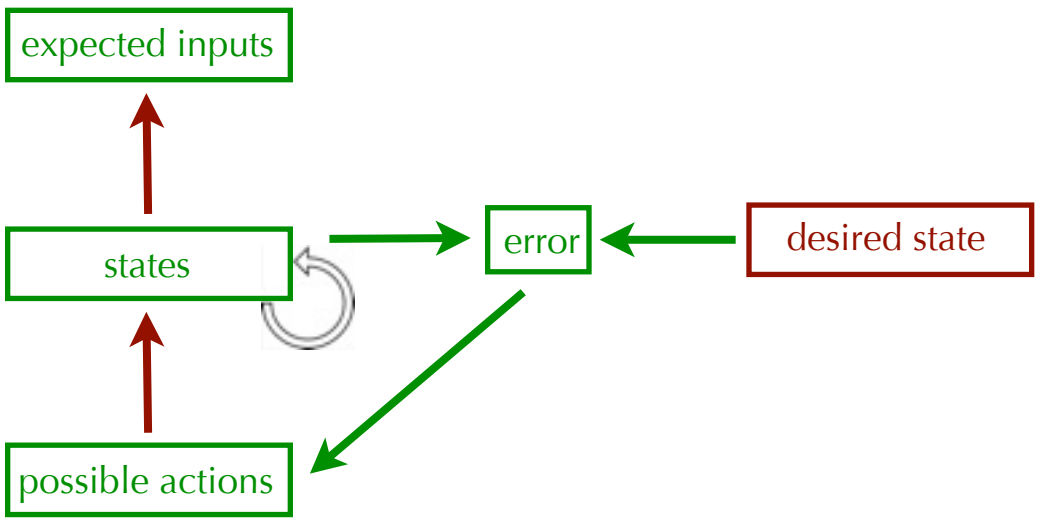
- Fixing the learned forward model to search for actions to achieve a perceptually-specified goal



- Key: Frozen/fixed by the world; Variable (error minimisation)

Planning as error-minimisation

- Fixing the learned forward model to search for actions to achieve an abstractly-specified goal

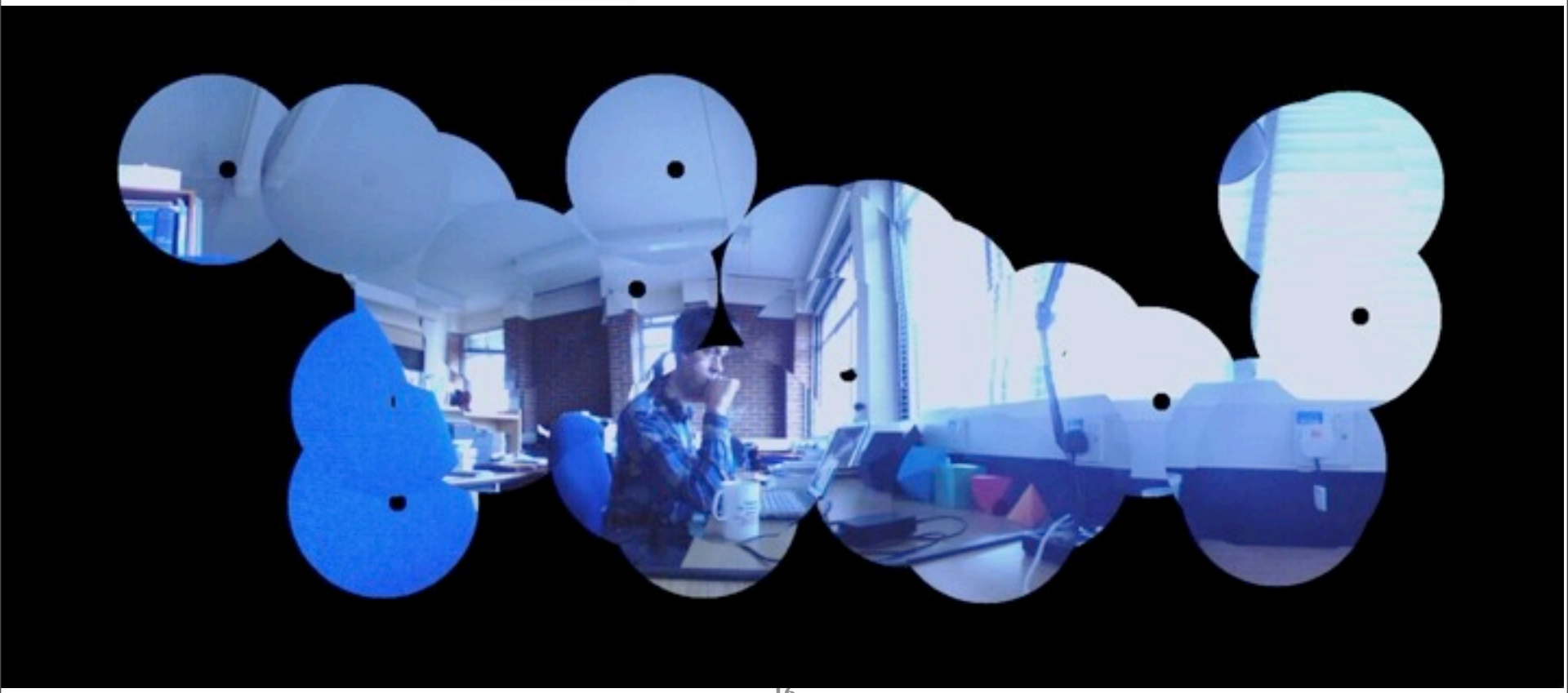


- Key: Frozen/fixed by the world; Variable (error minimisation)

An expectation-based theory of consciousness

- A discriminative, not reductive, theory of consciousness
  - Not trying to solve the hard problem, but rather answer the question: Assuming a given system is conscious, which conscious state is it in? (what is the content of its experience?)
- Suppose:
  - C = part of the non-conceptual content of the visual experience of a subject at a time
  - EBA contains E, a forward model that maps possible actions to expected foveal inputs
  - E(a) = the foveal input EBA's forward model would expect to receive were it to perform a in the current context
- Then the theory claims that, roughly put:
  - C = the conjunction, for all actions a, of E(a) with each E(a) located in the visual field in a way isomorphic to the spatial relations between the actions a

A depiction of the EBA's expectational state



The promise of EM

- EM was originally constructed to assist in the precise specification of experiential content ("synthetic phenomenology")
- (It also seems to me that) a model like EM might be able to account for the phenomenology of, e.g.:
  - Change blindness
  - Inattentional blindness
  - Troxler fading
  - Filling-in
  - After-images
  - Neglect
  - Eye-position contingent perception

A conceptual, not empirical contribution

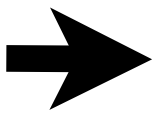
- But the preceding slide is largely speculation/hand-waving:
  - I am not a vision scientist; I am presenting no data today
- A reminder of what I am doing:
  - a) identifying the challenges that confront PC models of consciousness
  - b) suggesting some ways to overcome them
  - c) identifying which features of PC models facilitate this and which do not

From predictive coding models to an expectation-based model

Features retained, transformed, demoted

Typical PC model

- Predictive
- Error-minimising
- Inferential (Helmholtz)
- Action-involving
- Hierarchical (priors)
- Probabilistic
- Optimal (Bayesian)
- Subtractive ("suppressive")



Expectational Model

- **Predictive**
- **Error-minimising**
- Holistic (Merleau-Ponty)
- Enactive
- Hierarchical (features)
- Probabilistic
- Optimal (Bayesian)
- Subtractive ("suppressive")

From PC to EM: Retained

- **Predictive**
  - Although of a particular (enactive) sort (see "Enactive" below)
- **Error-minimising**
  - Yes, even extending to action, but in a way that supports/explains basic means/end rationality, not undercutting it (see "Enactive" below)

From PC to EM: Transformed

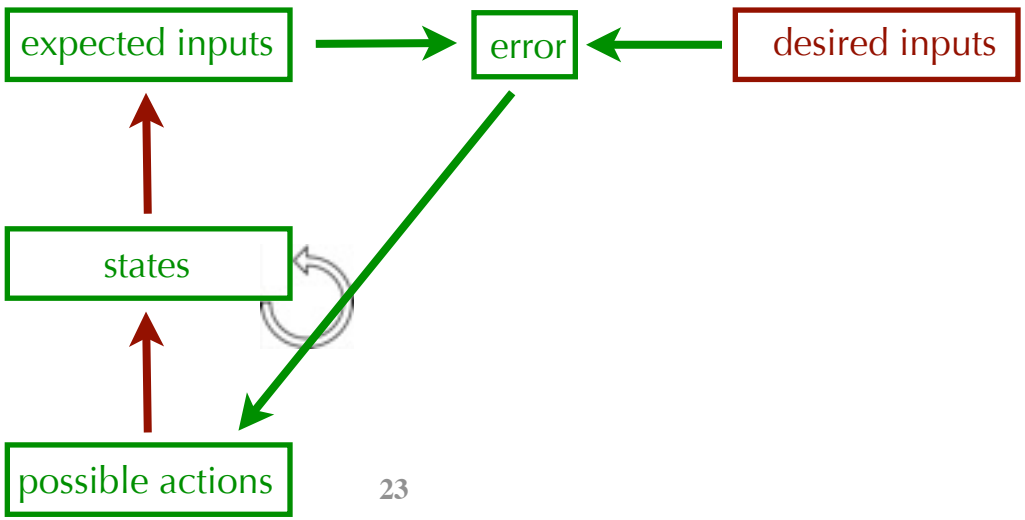
- **From Helmholtz's inference to Merleau-Ponty's holism**
  - "Inference": too cognitive/linguistic/conceptual?
  - Issue at stake: M-P's rejection of the "constancy hypothesis" -- the content of experience cannot be "read off" the sensory surface at a time

From PC to EM: Transformed

- **From action-involving to enactive**
  - What is fundamental is predictions of a particular sort: expectations of how the world will change if I act this way or that.
    - At root, not detached, disembodied abstract causal reasoning
    - Relations between actions actually structure the phenomenal space
  - Action is incorporated in a way that supports/explains basic means/end rationality, not undercutting it
    - Friston: As for model selection so also for action selection
    - We do not choose the action that achieves some goal, but rather we perform the action that minimises prediction error
    - Problems:
      - Darkroom (Firston, Thornton and Clark 2012)
      - Radical undermining of what it is to be a cognitive, intentional agent

From PC to EM: Transformed

- EM *does* find a use for the notion of action-selection that minimises error
- But not, like Friston, with respect to *prediction* error (*actual* world)
- Rather, *expected* error (*desired, non-actual* world)
- Cf third, "planning" deployment of the forward model in EBA:

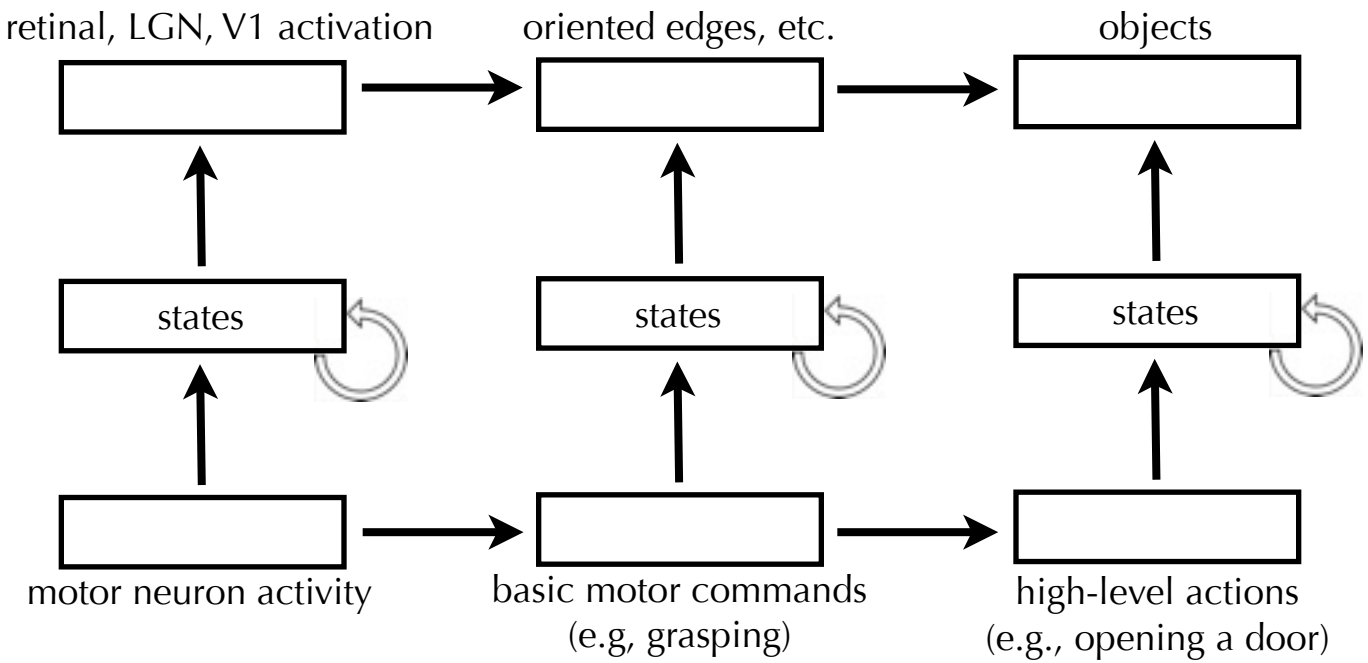


23

From PC to EM: Transformed

- **Hierarchy of priors to hierarchy of features**
  - One of the strengths of the PC approach is its extendability to higher forms of perception and cognition
  - Achieved by cascading models at different levels of abstraction so that the priors assumed in a lower level are fixed by the level above it.
  - A similar, but distinct (given its non-Bayesian structure) kind of hierarchy can be constructed for EBA, and thus EM

Hierarchical extension



- Higher level expectational forward models, grounded in lowest

Hierarchical extension

- Higher-level expectational forward models
- Grounded in lowest level (source of error signal)
- Allows the formulation of a range of hypotheses extending EM
  - What role do these higher-level expectations play in determining experiential content?
  - Do they only play a role in determining the content of experience in so far as they have a top-down impact on the lowest level (left hand side) of expectations (which alone directly determine the content of experience)?
  - Do they constitute higher layers of phenomenal content, that are experienced in conjunction with the content determined by the lower layer?
  - Both?
  - Neither?

From PC to EM: Demoted

- EM is explicated without essential reference to:
  - Expectations as **probabilistic**
  - The settling on one set of expectations as a (Bayesian) **optimal** process
  - The top-down **suppression** of expected components of the input signal
- These features are consistent with EM, and thus could be added if needed
  - E.g., it seems likely that attentional sub-systems would have use for a reduced, error-enhanced signal

From PC to EM: Demoted

- On the other hand, some of these features of PC models may impede progress on the kind of account of consciousness that EM aims to provide
  - Probabilistic but Optimal:
    - On the one hand, as (Clark 2012) points out, there is a disconnect between the probability and the apparent unity and determinateness of phenomenal experience
    - Of course, some (e.g. Dennett 1991) would reject the determinateness of experience, choosing instead to see it as being relative to how it is probed. But this is not the kind of indeterminateness probabilistic PC models can explain.
    - EM, on the other hand, can handle both: determinateness; or superposition of determinate contents
  - Subtractive
    - Leaving this as an option, rather than a defining feature of the framework, permits accounts of a (less radical) phenomenology in which we *don't* only experience the unexpected

Further differences between EM and standard PC models

- EM is:
  - Robustly experiential
    - Offers an account of the aspectual shape of experience, not just what the experience is about (unlike, e.g., (Hohwy, Roepstorff & Friston, 2008)?)
  - Counterfactual
    - The content of experience is not given (only) by the prediction of what input *will* be received if the current action is carried out
    - But also in terms of the expectations of what inputs *would be* received if an entire range of non-actual actions were individually carried out

Robustly experiential

- What makes consciousness science distinct from other, related sciences is that it aims not just to provide an account of the processes that correlate with or give rise to consciousness, but of the "what it is like"-ness of experiences themselves.
- In terms of Frege's famous distinction, a robustly experiential account of consciousness would involve not only what an experience is of (its reference), but also they way that "referent" is experienced (the experience's sense, or content).
- EBA's ambitions are robustly experiential in that it aims to give an account of the fine-grained non-conceptual content of visual experience: the aspectual shape of the experience, not just what the experience is about

Further differences between EM and standard PC models

- EM is:
  - Non-monolithic
    - E.g., no insistence on universal application of error suppression
    - But also: no insistence even on universal error-minimisation
      - For example, creative/novelty-seeking forms of cognition might result from an error-minimising predictive model being locked in a spiralling "arms race" with a model that seeks out inputs that exceed (but only just!) the model's current predictive capabilities (the "subjective edge of chaos"; Chrisley 2008)

Further differences between EM and standard PC models

- EM is:
  - Embodied
    - E.g., the spatial structure of experience depends on actual spatial structure of actions, not representation of such
    - But also: which actions are to be included in the set of those which determine the experience-generating expectations at any time may depend on whether they would actually provide information about parts of the world from which one is receiving (perhaps impoverished) visual information -- not representation of such
- This is not to say that no other PC models have these features
  - E.g., (Seth, Critchley and Suzuki 2012) presents an account that is robustly experiential and enactive (and affective)


Thank you.
Comments welcome:
ronc@sussex.ac.uk
33