

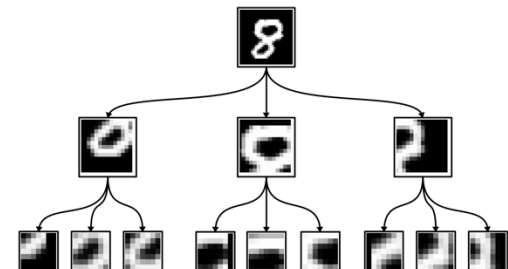
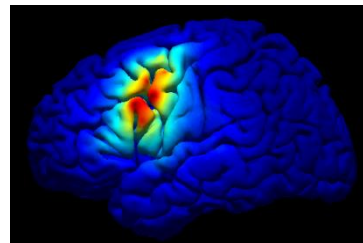
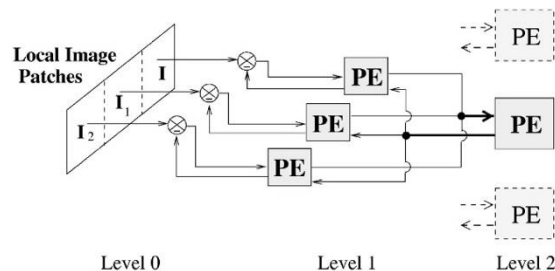


Predictive Coding

Rajesh P. N. Rao

CJ and Elizabeth Hwang Endowed Professor of CSE and ECE
Co-Director, Center for Neurotechnology
University of Washington, Seattle, USA

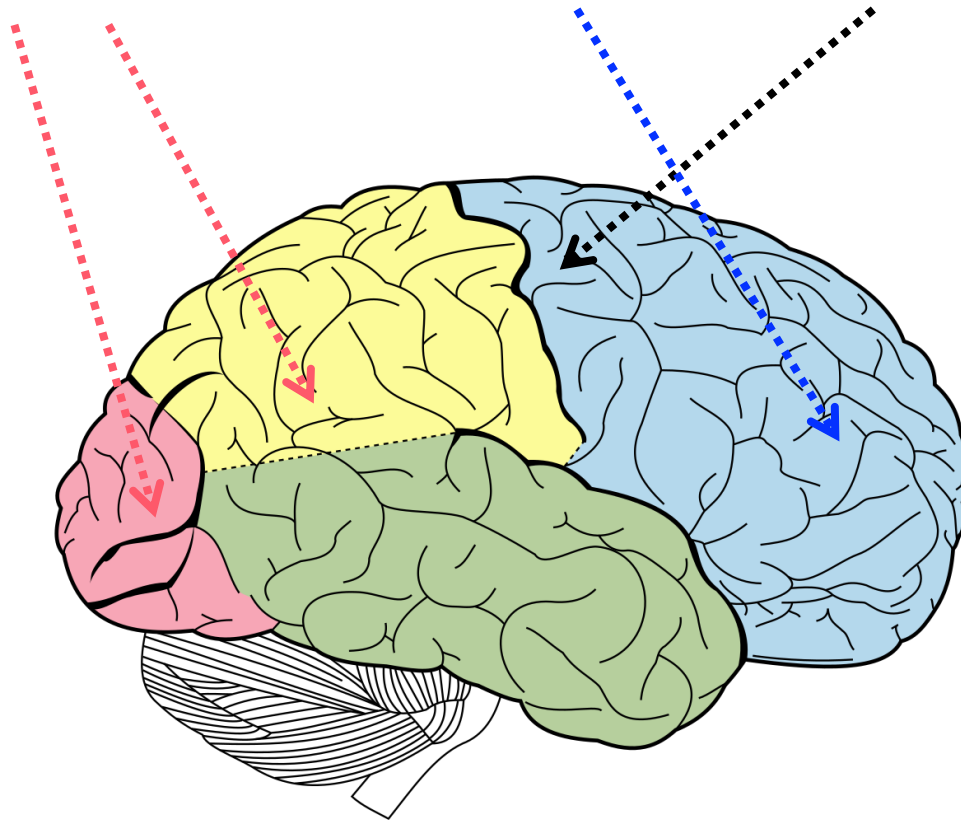
October 9 2026



Traditional Model of Intelligent Systems

“Perception” → *“Cognition”* → *“Action”*

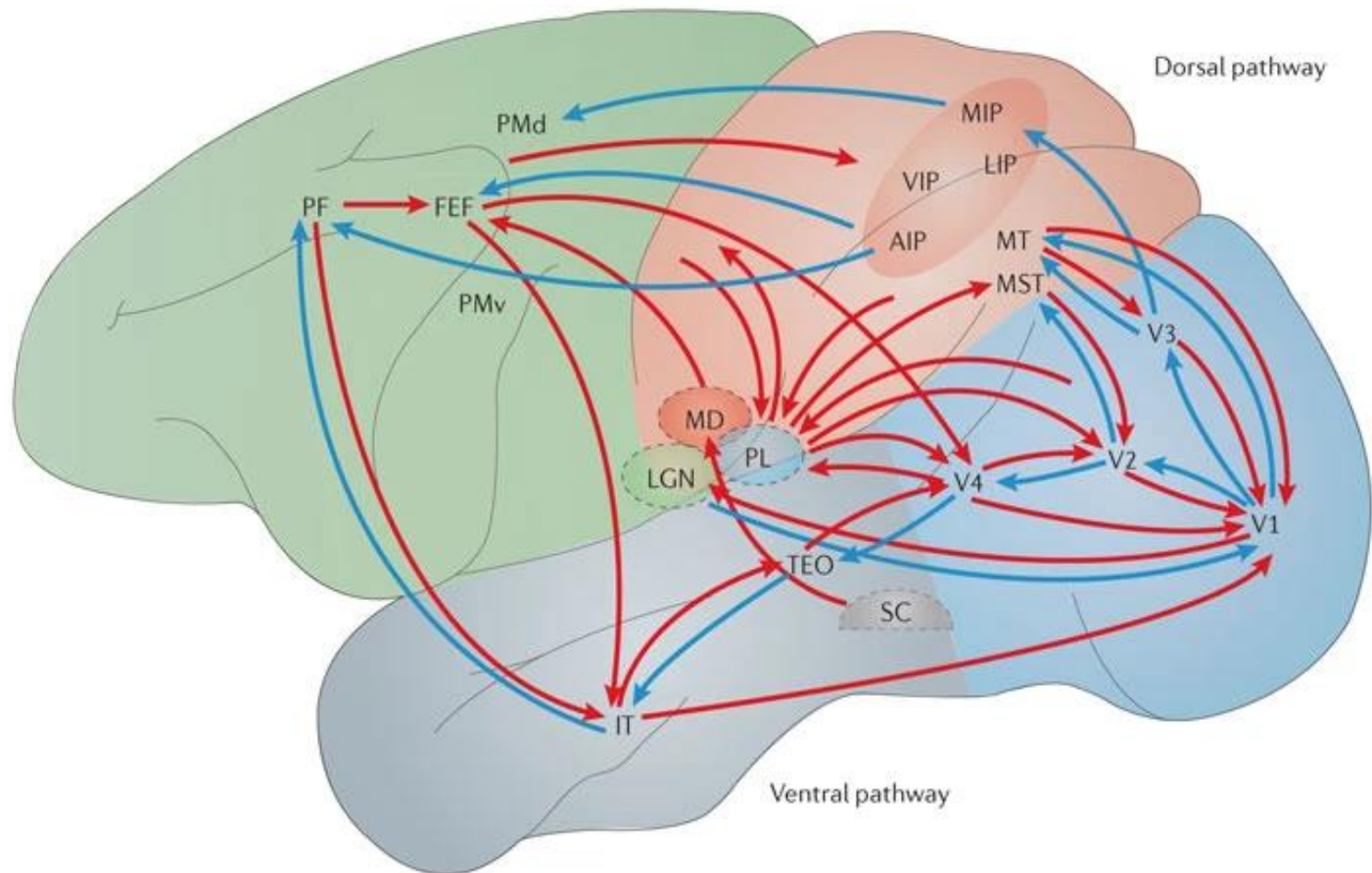
Process Input → *Make Decision* → *Move and act*



The Brain is NOT a Feedforward Network

Feedforward: Blue

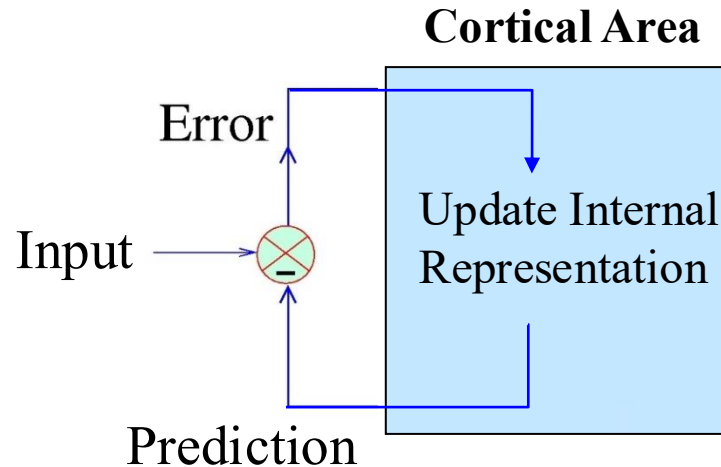
Feedback: Red



(Gilbert & Li, 2013)

Nature Reviews | Neuroscience

Predictive Coding Theory of Cortex (1999)



(Rao & Ballard, 1997,1999; Rao, 1999; Huang & Rao, 2011; Jiang & Rao, 2021)

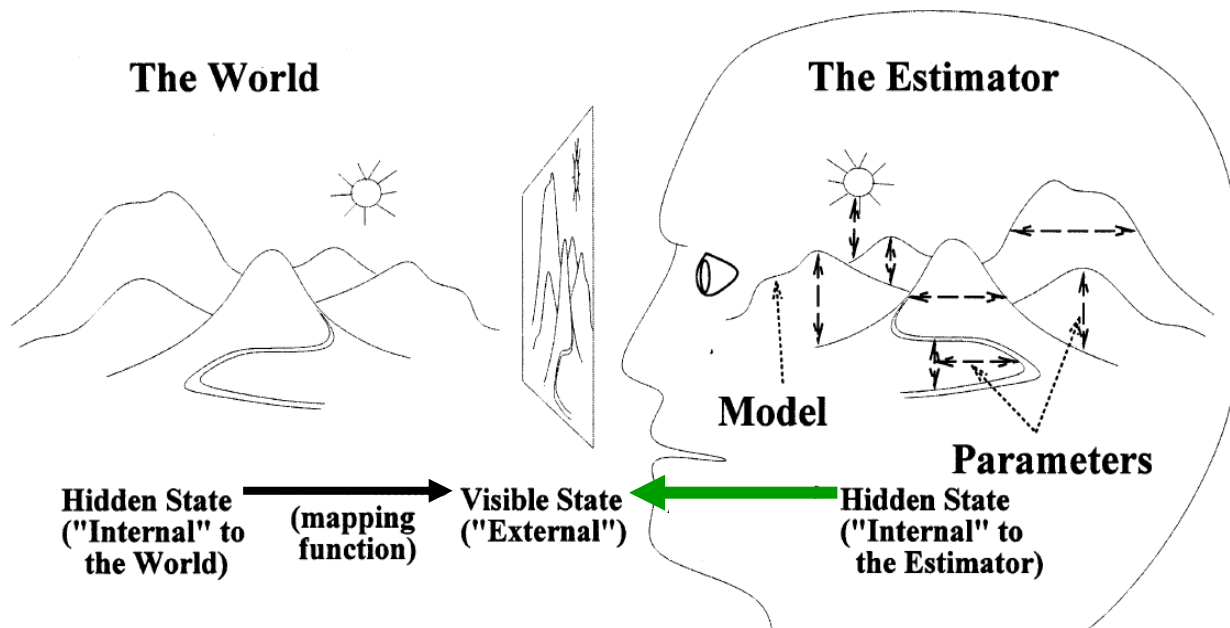
Feedback = Prediction

Feedforward = Prediction Error

Implements Bayesian Inference

Bayesian Brain Hypothesis

- Hypothesis: The neocortex learns an **internal (generative) model of the world** and uses **Bayesian inference** to estimate the most probable world states causing sensory inputs

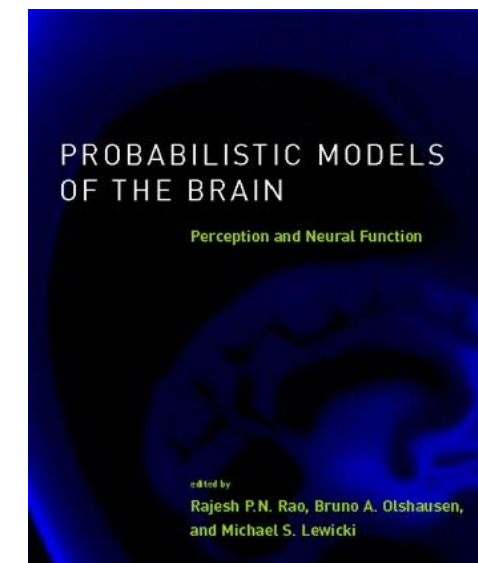


Bayesian Brain

PROBABILISTIC APPROACHES
TO NEURAL CODING



edited by
KENJI DOYA, SHIN ISHII,
ALEXANDRE POUGET,
AND RAJESH P. N. RAO



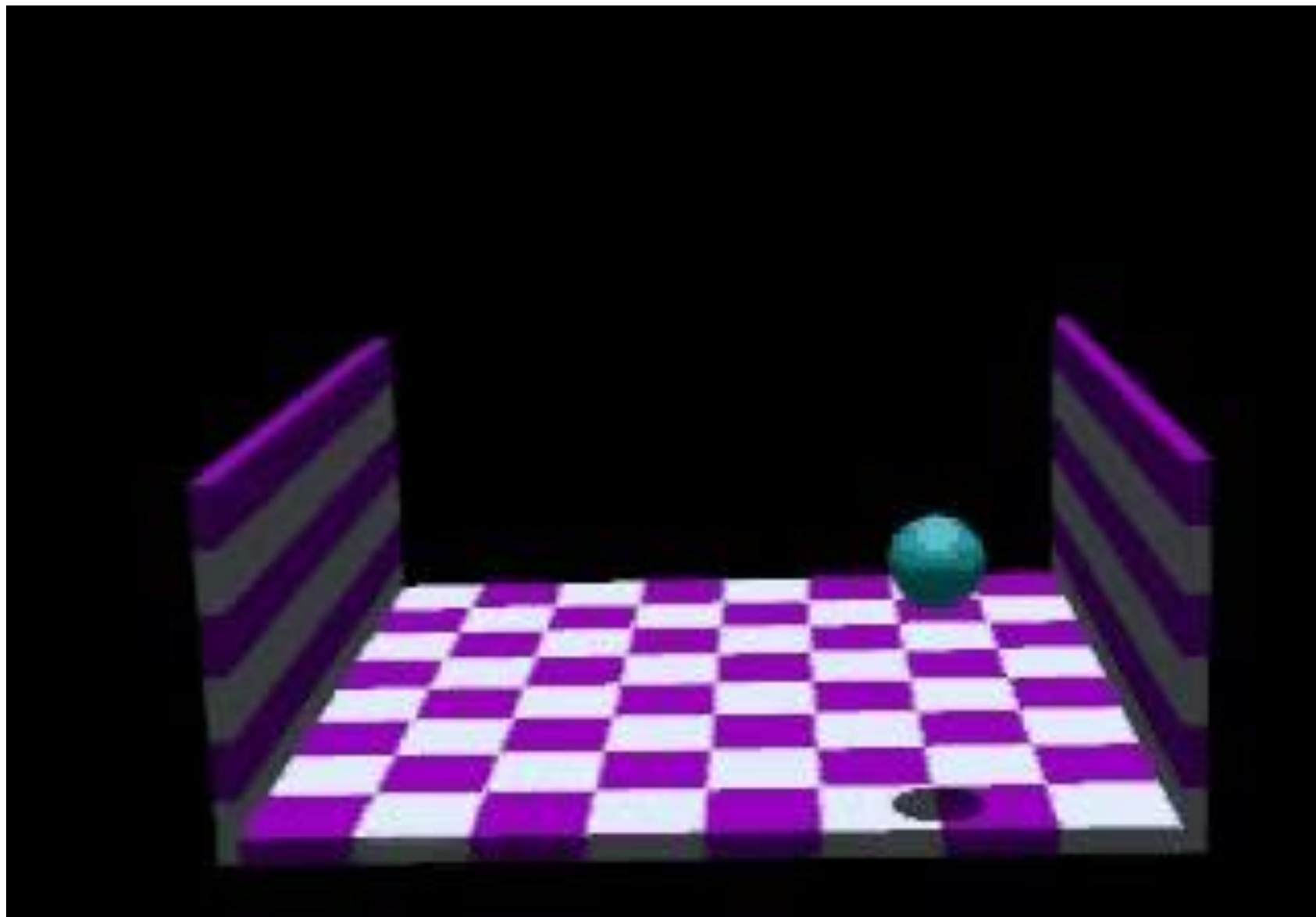
PROBABILISTIC MODELS
OF THE BRAIN

Perception and Neural Function

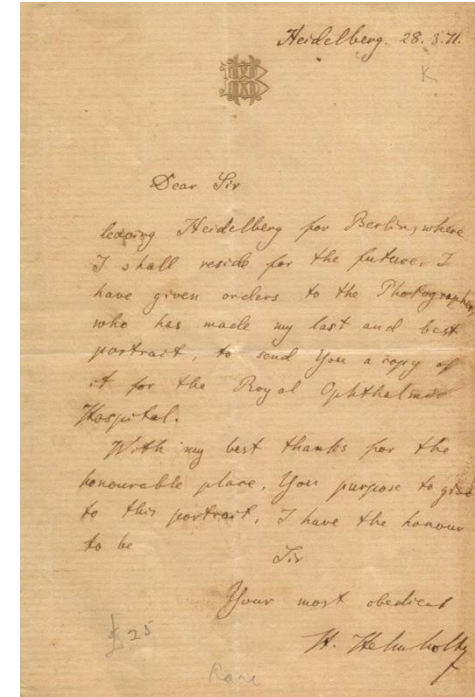
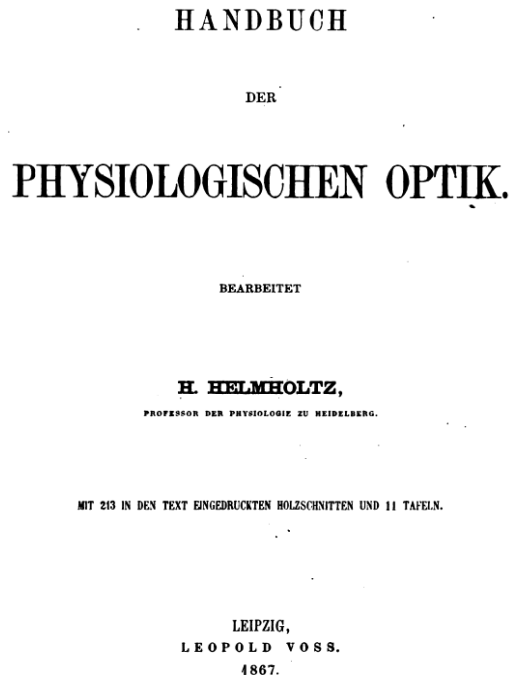
edited by
Rajesh P.N. Rao, Bruno A. Olshausen,
and Michael S. Lewicki

Experimental Evidence for Bayesian Models

- ❑ Motion from cast shadows (Kersten et al., 1996)
- ❑ Surface perception based on texture (Knill, 1998)
- ❑ Inferring 3D shape from 2D images (Mamassian et al., 2002)
- ❑ Color perception (Bloj et al., 1999)
- ❑ Cue combination for depth perception (Jacobs, 2002)
- ❑ Motion illusions (Weiss et al., 2002)
- ❑ Motor control (Körding and Wolpert, 2004)
- ❑ Speed perception (Stocker and Simoncelli, 2006)
- ❑ and many others....



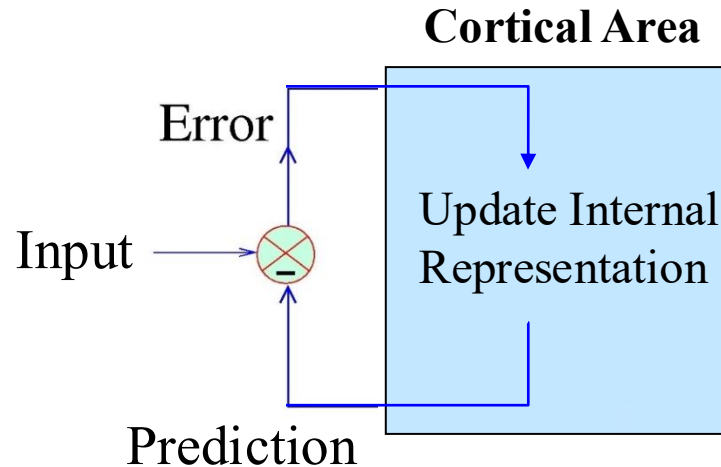
Origins: Helmholtz



Perceptual processes “...are in general not conscious but **unconscious**. In their outcomes they are like **inferences** insofar as we from the observed effect on our senses arrive at **an idea of the cause of this effect.**”

Helmholtz, 1867; page 430

Predictive Coding Model: An Implementation of Bayesian Inference

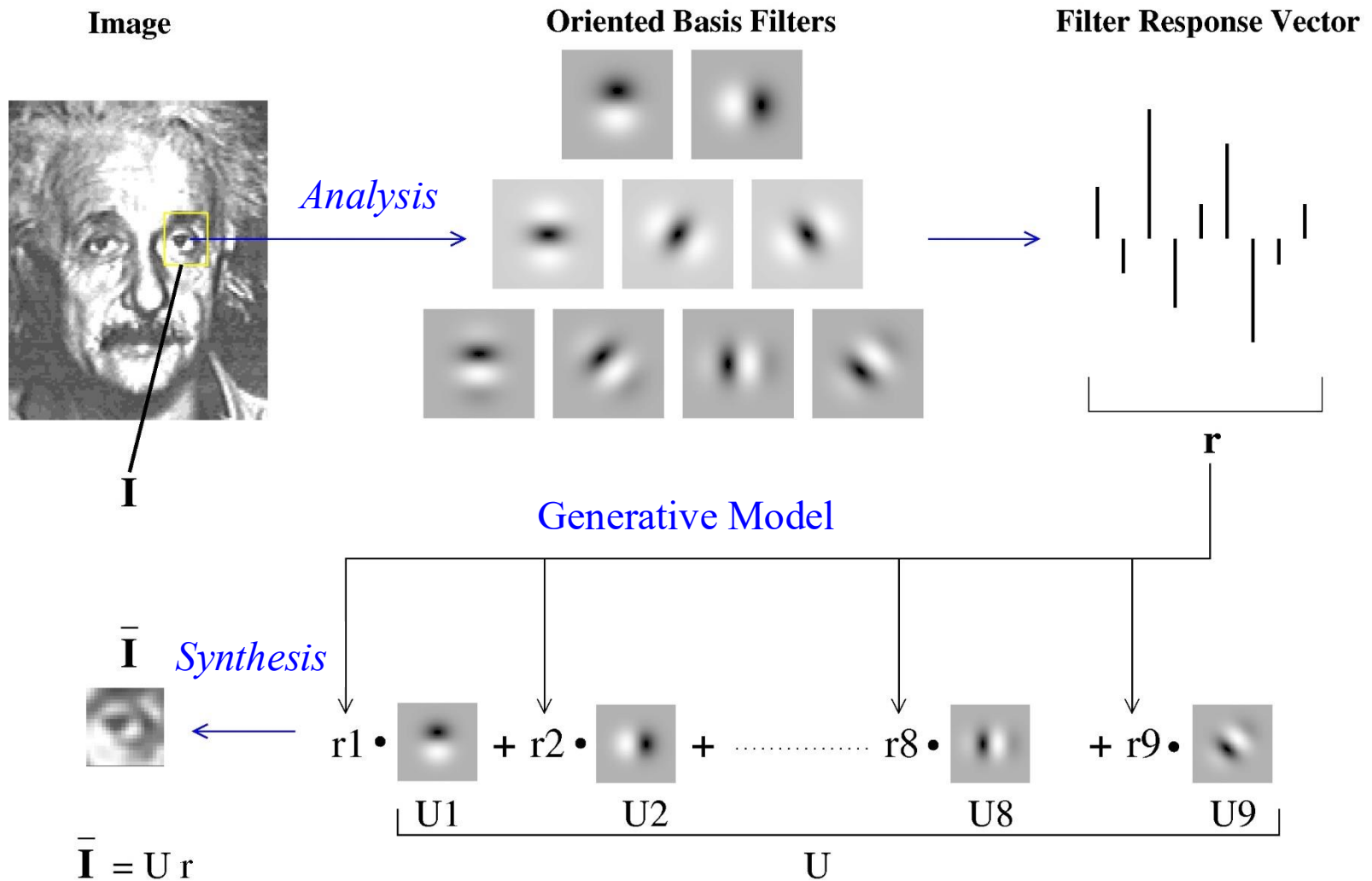


(Rao & Ballard, 1997,1999; Rao, 1999; Huang & Rao, 2011; Jiang & Rao, 2021)

Cortical Feedback = Top-Down Prediction
Feedforward Connections = Prediction Error

Based on a *generative model* of inputs

Simple Example: Linear Generative Model



Spatiotemporal Generative Model

▣ Spatial Generative Model:

$$\mathbf{I}(t) = U\mathbf{r}(t) + \mathbf{n}(t)$$

$\mathbf{r}(t)$ = hidden state vector, $\mathbf{n}$ = zero mean Gaussian white noise with covariance Σ

▣ Temporal Dynamics of State Vector:

$$\mathbf{r}(t) = V\mathbf{r}(t-1) + \mathbf{m}(t-1)$$

V = transition matrix, $\mathbf{m}$ = zero mean Gaussian white noise with covariance Σ_m

▣ **Goal:** Minimize Prediction Errors (weighted by *inverse variance* aka *precision*):

$$E = (\mathbf{I} - U\mathbf{r})^T \Sigma^{-1} (\mathbf{I} - U\mathbf{r}) + (\mathbf{r} - \bar{\mathbf{r}})^T M^{-1} (\mathbf{r} - \bar{\mathbf{r}})$$

Minimize w.r.t. $\mathbf{r}$ = Inference via Network Dynamics

Minimize w.r.t. Synaptic Weights U, V = Plasticity/Learning

Minimize w.r.t. Actions = Planning/Active Inference

Minimizing Prediction Errors via State Estimation

- Setting $dE/d\mathbf{r} = 0$ and solving for the optimal $\mathbf{r}$ yields **network dynamics for minimizing prediction errors**:

$\bar{\mathbf{r}}(t) = V \hat{\mathbf{r}}(t-1)$ *Next state prediction (lateral connections)*

$\bar{\mathbf{I}} = U \bar{\mathbf{r}}(t)$ *Top-down feedback prediction of input*

$\hat{\mathbf{r}} = \bar{\mathbf{r}} + \underbrace{G(t)}_{\text{Gain}} \underbrace{U^T(\mathbf{I}(t) - U \bar{\mathbf{r}}(t))}_{\text{Error}}$ *Correction based on error*

U^T = Feedforward Matrix

U = Feedback Matrix

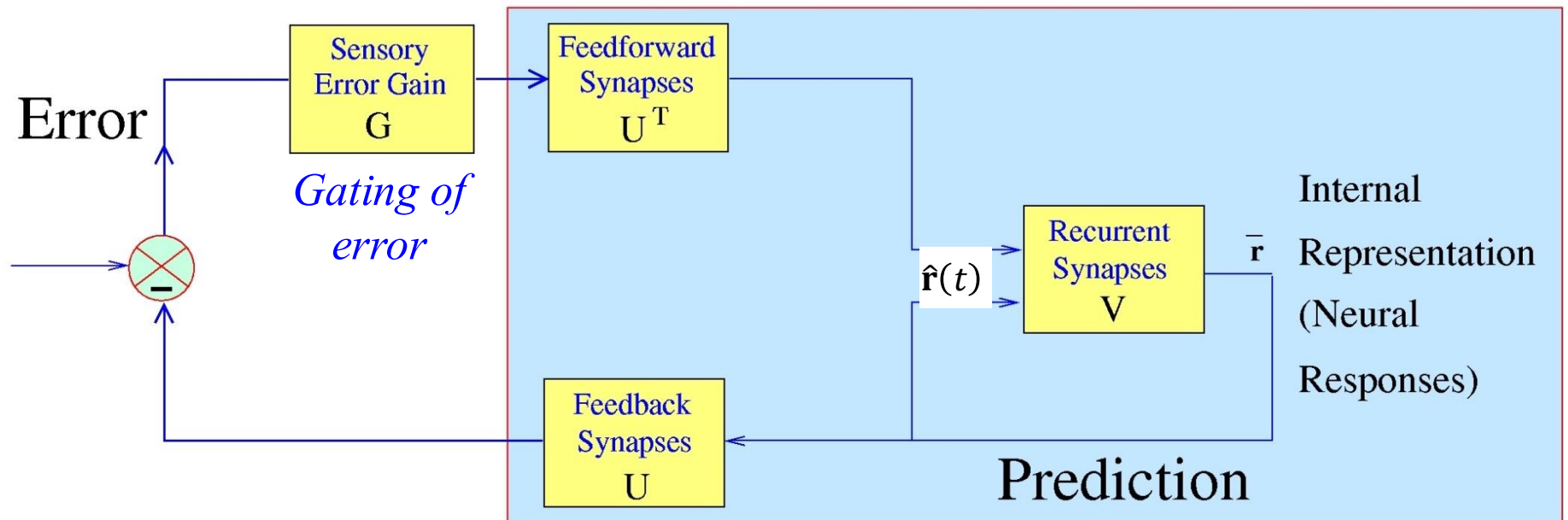
V = Recurrent Matrix (Lateral Connections)

G = Sensory Error Gain / Precision Matrix

Single Level Predictive Coding Model

$$\hat{\mathbf{r}}(t) = \bar{\mathbf{r}}(t) + U^T G(t)(\mathbf{I} - U\bar{\mathbf{r}}(t)) \quad \text{Prediction} \quad \bar{\mathbf{r}}(t) = V\hat{\mathbf{r}}(t - 1)$$

(Variance is updated as well)



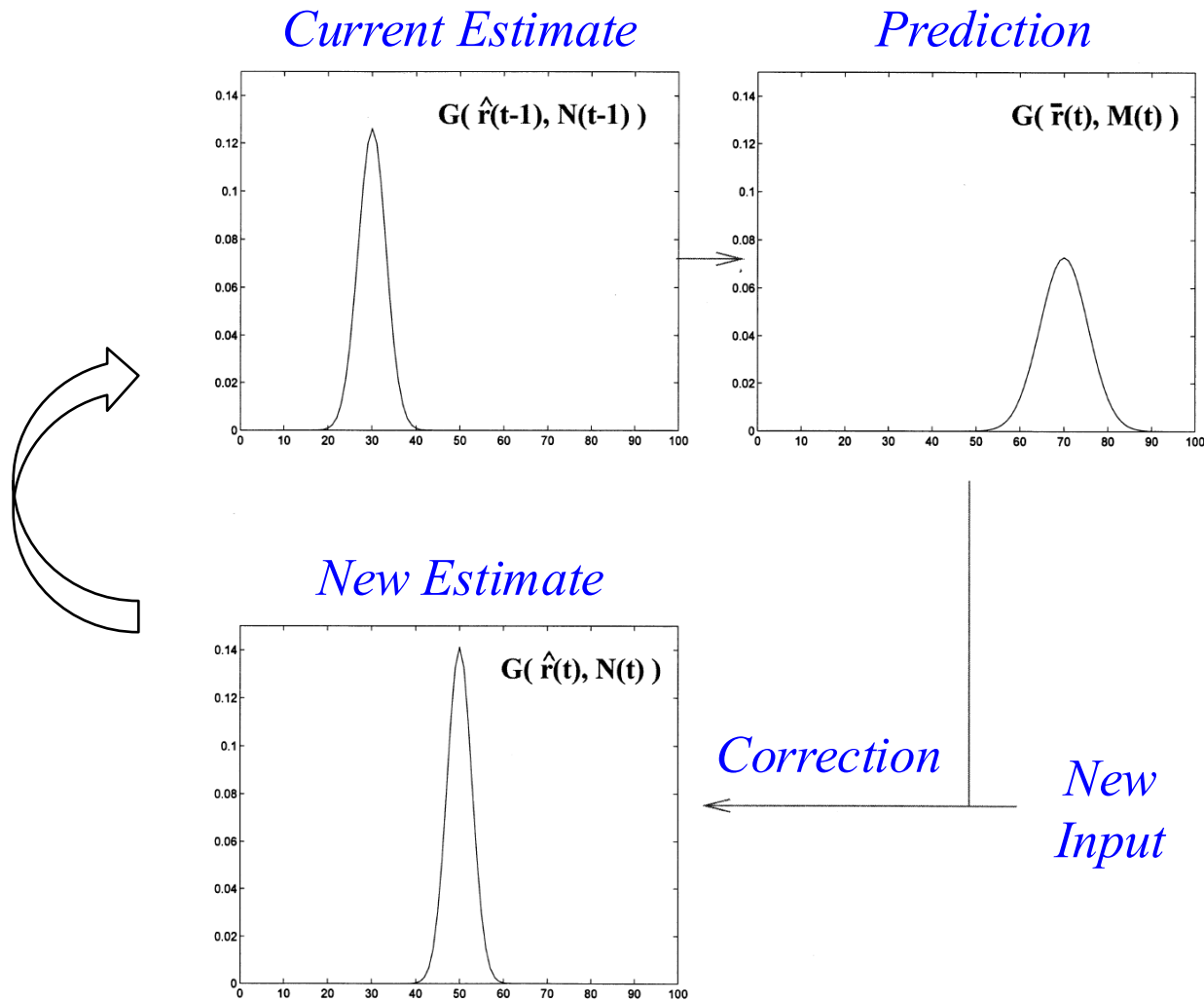
(Rao & Ballard, *Neural Comp* 1997, *Nature Neuro* 1999;
Rao, *Vis Res* 1999; Huang & Rao, *Wiley Cog Sci* 2011)

Cortical Feedback = Top-Down Prediction

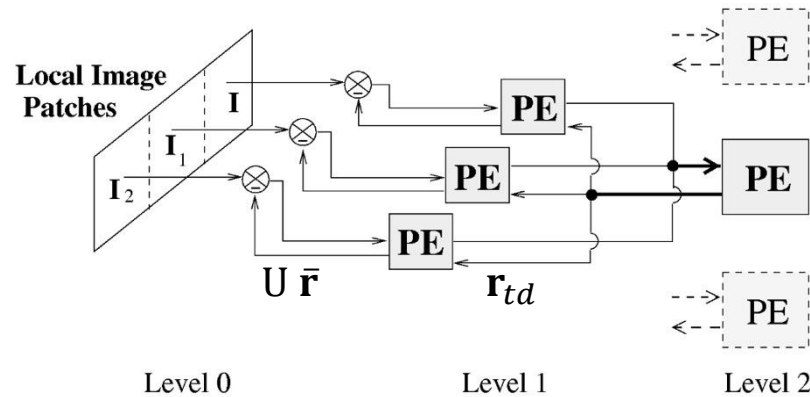
Feedforward Connections = Prediction Error

Local Recurrent Connections = Temporal State Prediction

Bayesian Inference via Prediction and Error Correction



Hierarchical Predictive Coding

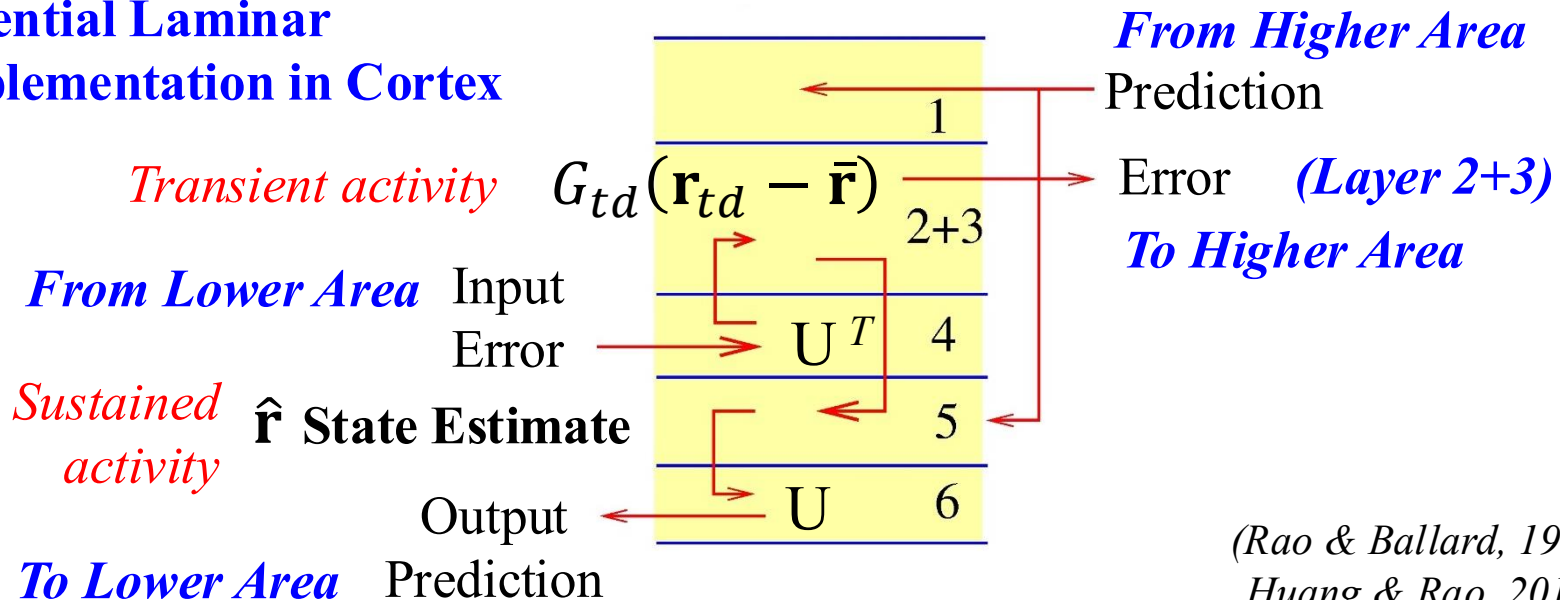


Top-down modulation



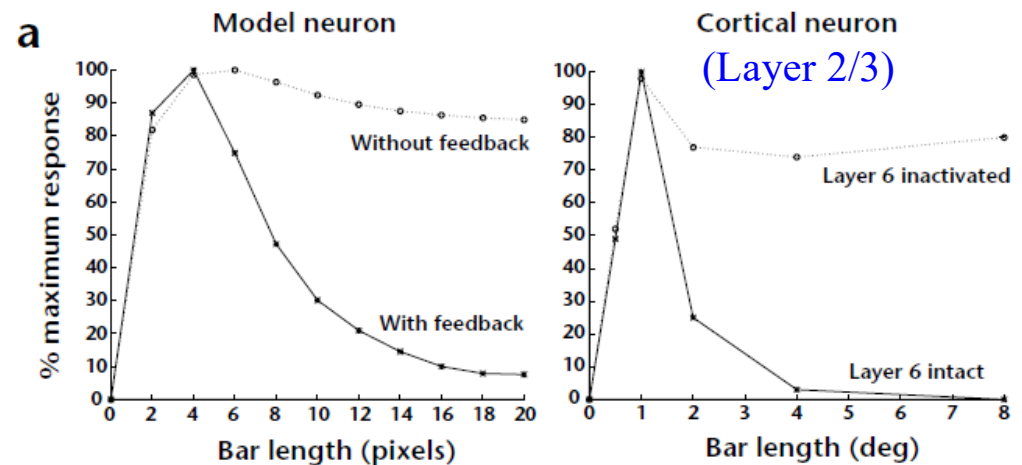
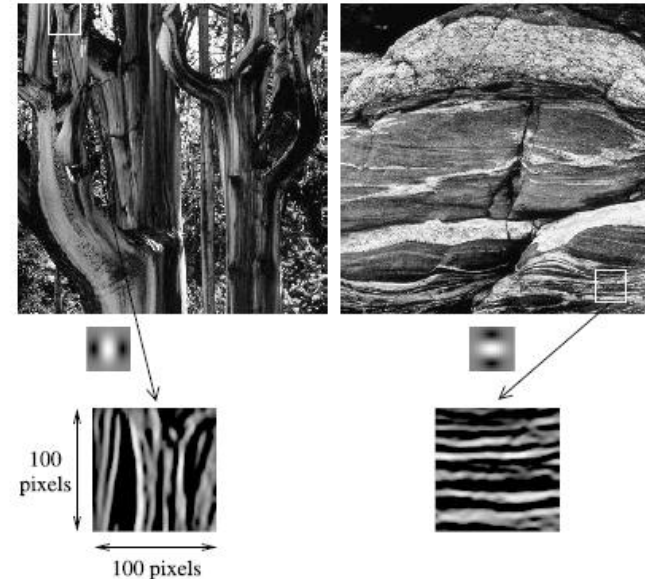
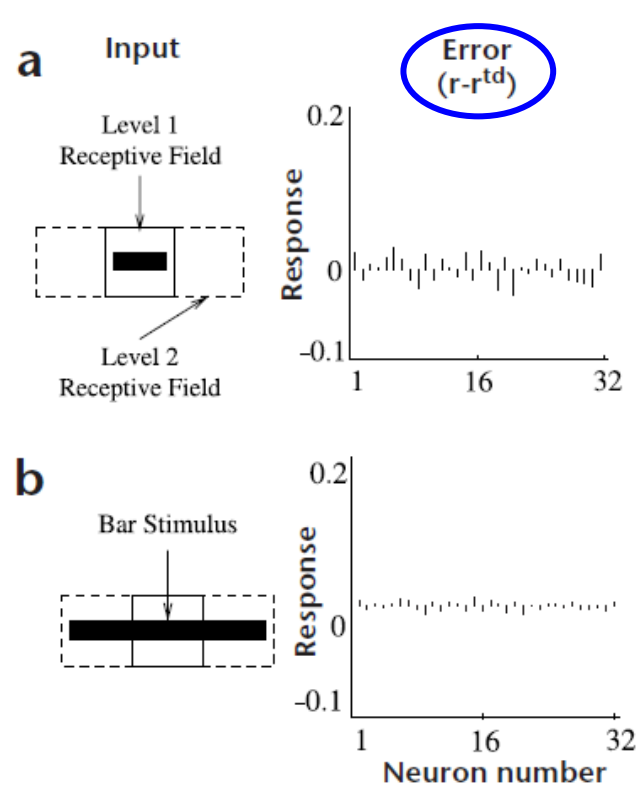
$$\hat{\mathbf{r}} = \bar{\mathbf{r}} + U^T G(\mathbf{I} - U \bar{\mathbf{r}}) + G_{td}(\mathbf{r}_{td} - \bar{\mathbf{r}})$$

Potential Laminar Implementation in Cortex



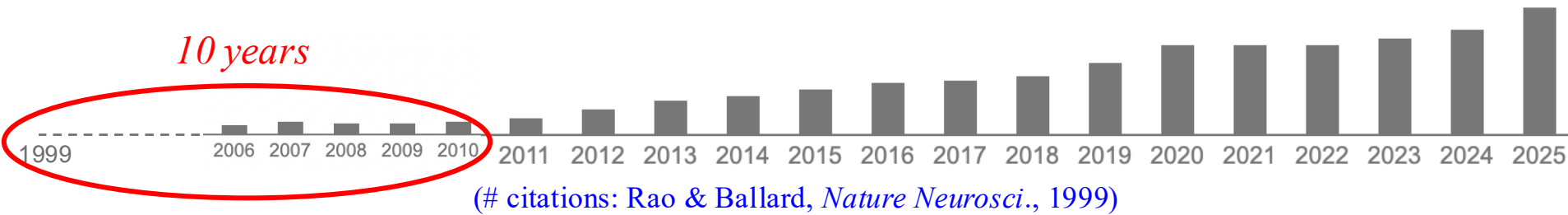
(Rao & Ballard, 1999;
Huang & Rao, 2011)

Predictive Coding explains Contextual Effects

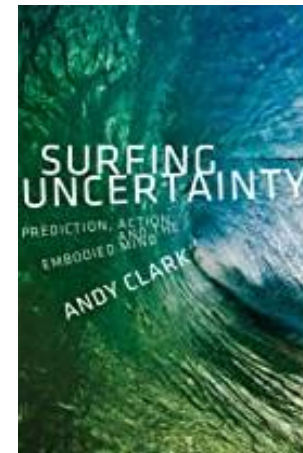
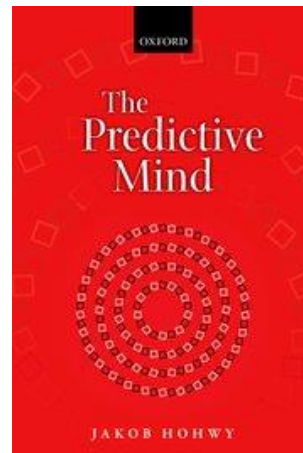
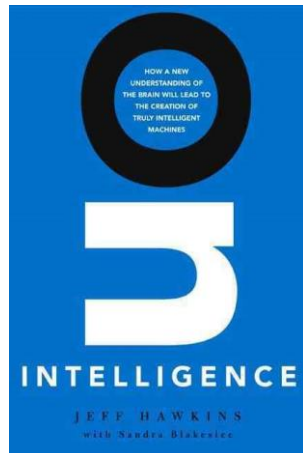


(Rao & Ballard, *Nature Neuroscience*, 1999)

Growing Interest in Predictive Coding in Recent Years



Attention (Feldman & Friston, 2010), binocular rivalry (Hohwy et al., 2008), neural repetition suppression (Summerfield et al., 2008), mirror neurons (Kilner et al., 2007), motor planning (Brown et al., 2011), action understanding (Ouden et al., 2012), theory of mind (Koster-Hale & Saxe, 2013), emotions (Seth, 2013), aesthetics (de Cruys & Wagemans, 2011), self-awareness (Apps and Tsakiris, 2014), consciousness (Seth et al., 2011), autism (Lawson et al., 2014), locomotion (Fridman & Petreanu, 2017), schizophrenia (Sterzer et al., 2018), automobile driving (Engström et al., 2018)...



Predictive Coding and Friston's Free Energy Principle



Variational inference: Approximate probability distribution p_θ with tractable distribution q_φ (parameters are θ, φ)

$$KL(q_\varphi(\mathbf{r}|\mathbf{I})||p_\theta(\mathbf{r}|\mathbf{I})) = \int q_\varphi(\mathbf{r}|\mathbf{I}) \log \frac{q_\varphi(\mathbf{r}|\mathbf{I})}{p_\theta(\mathbf{r}|\mathbf{I})} d\mathbf{r} = \int q_\varphi(\mathbf{r}|\mathbf{I}) \log \frac{q_\varphi(\mathbf{r}|\mathbf{I})}{p_\theta(\mathbf{r}, \mathbf{I})} d\mathbf{r} + \log p_\theta(\mathbf{I}) \\ = \mathcal{F} + \log p_\theta(\mathbf{I})$$

$$\mathcal{F} = KL(q_\varphi(\mathbf{r} | \mathbf{I}) || p_\theta(\mathbf{r} | \mathbf{I})) - \log p_\theta(\mathbf{I})$$

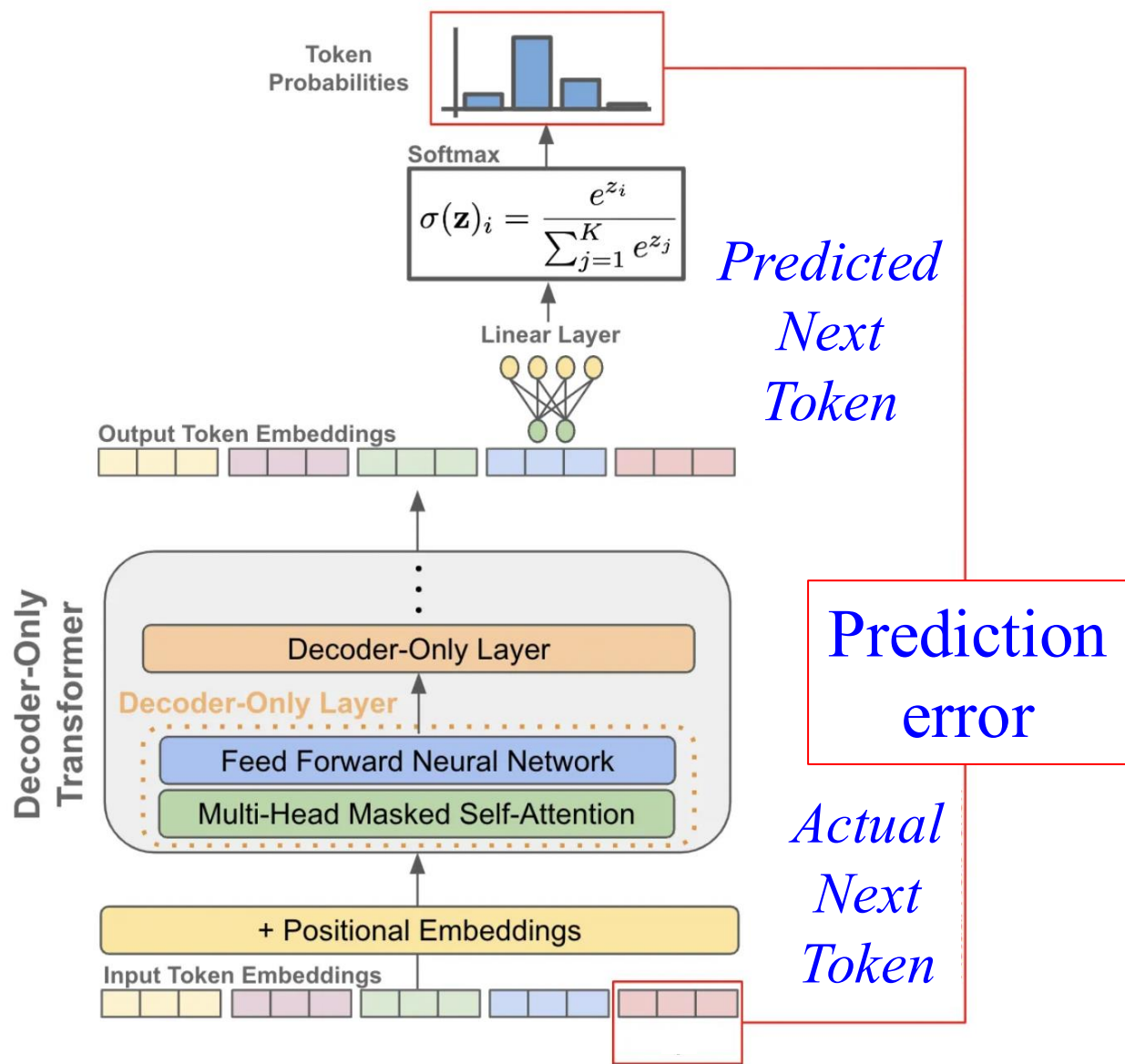
Free energy principle (Friston, 2010): Organisms minimize variational free energy $\mathcal{F}$ to minimize KL divergence between q and p . Can rewrite $\mathcal{F}$ for predictive coding as:

$$\mathcal{F} \propto KL(q_\varphi(\mathbf{r} | \mathbf{I})||p_\theta(\mathbf{r})) + E_q \left[\frac{1}{\sigma^2} \|\mathbf{I} - f(\mathbf{U}\mathbf{r})\|^2 \right] \quad (\text{cf. VAE loss function})$$

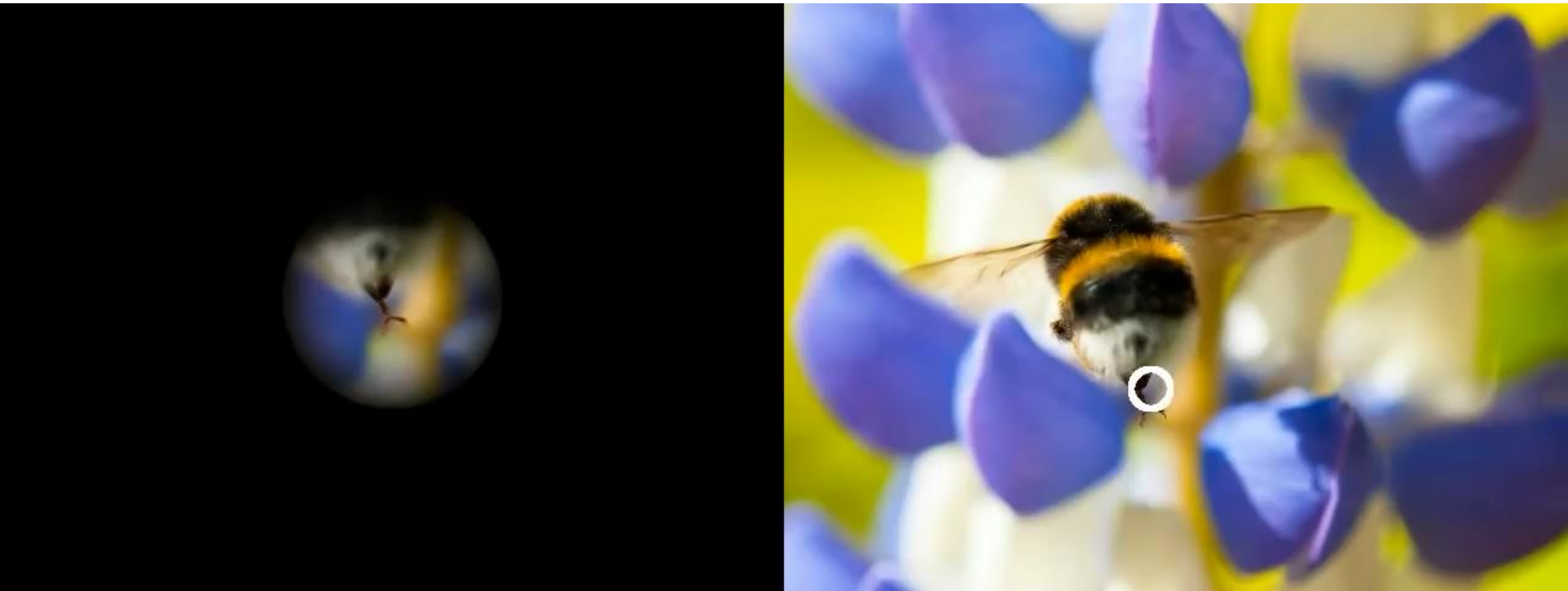
= KL-divergence between posterior and prior for $\mathbf{r}$ + Mean Squared Prediction Error

Minimizing predictions errors (plus KL divergence from prior) minimizes variational free energy

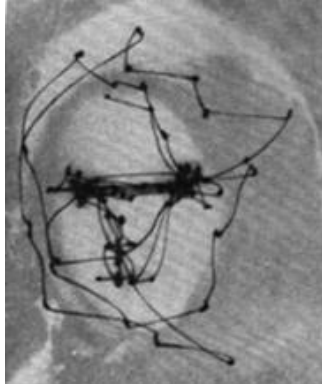
2018: AI discovers Prediction Errors for training Transformers



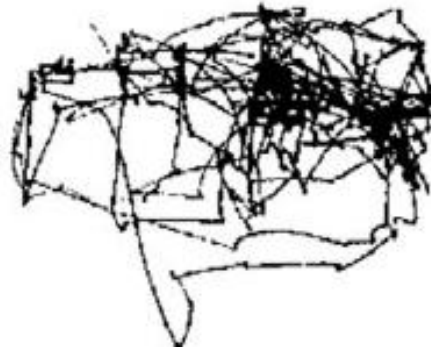
BUT...the predictive coding model
ignored an important aspect of vision
and all of perception...



Perception involves Action!

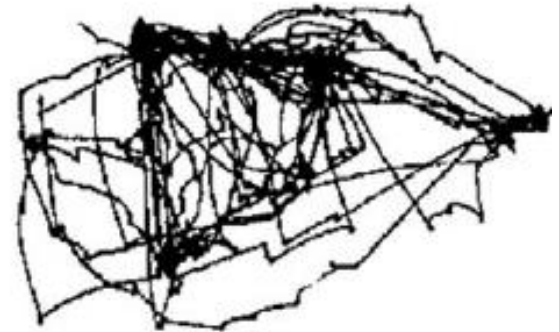


(Yarbus, 1967)



Estimate how wealthy they are

Eye movements
(Free exploration)



Estimate their ages

Perception is not static – it involves eye movements, head movements, body movements...

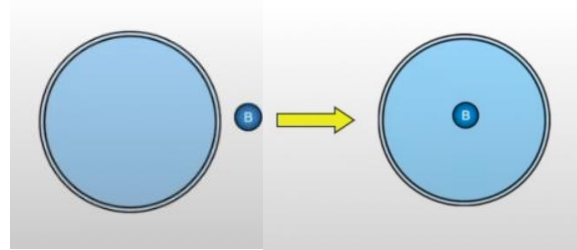
~~How~~ did “intelligence” evolve?

Why

When life was simple...

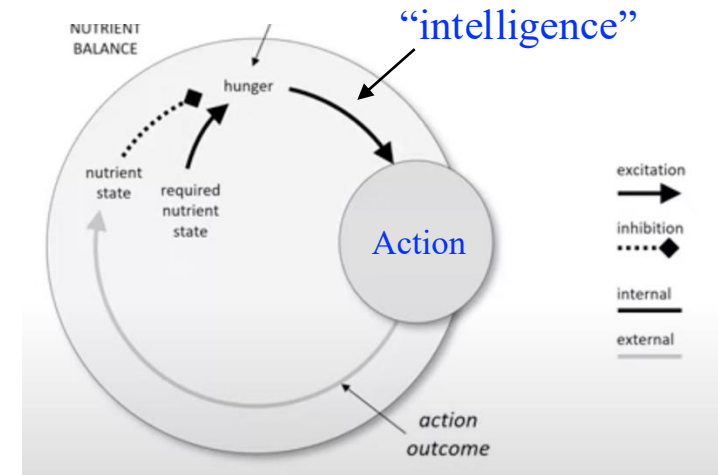


Later



*Internal
metabolic control
via chemistry*

*External control
of environment
via action*



(Cisek, 2019)

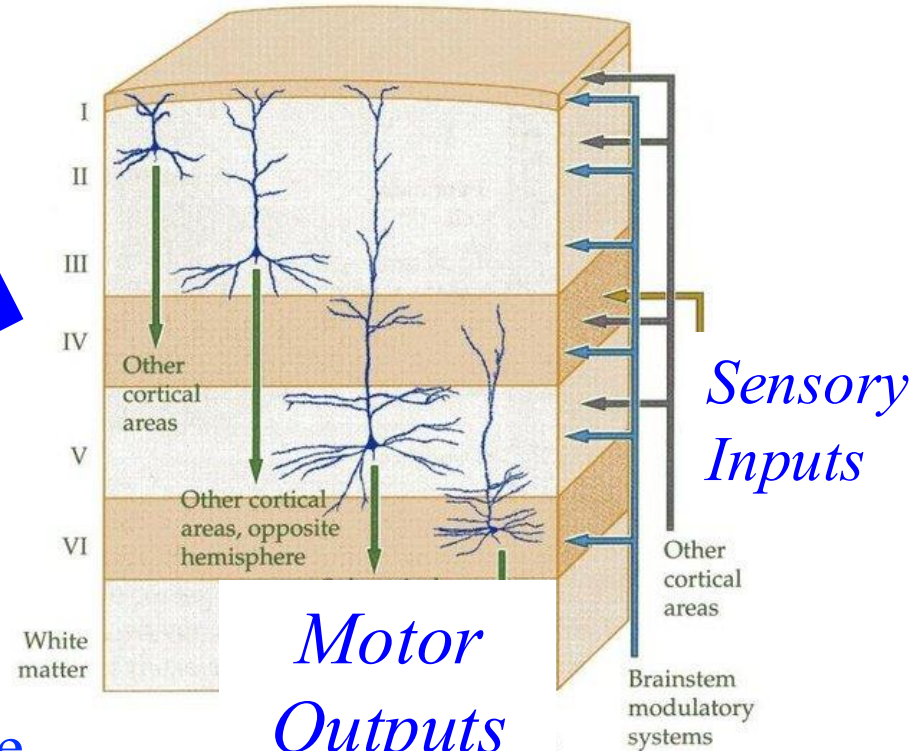
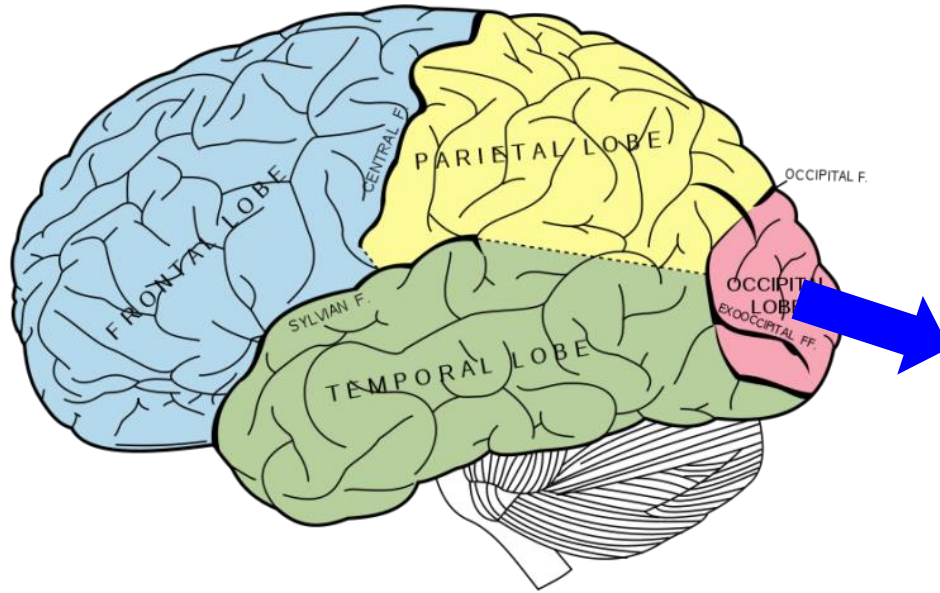
Related ideas: Buzsáki,
Llinás, Powers, Friston, ...

Intelligence likely evolved to control the body:

- Move towards desirable conditions (nutrients, food)
- Move away from undesirable conditions (noxious stimuli, predators)

(Examples adapted from Paul Cisek's talk)

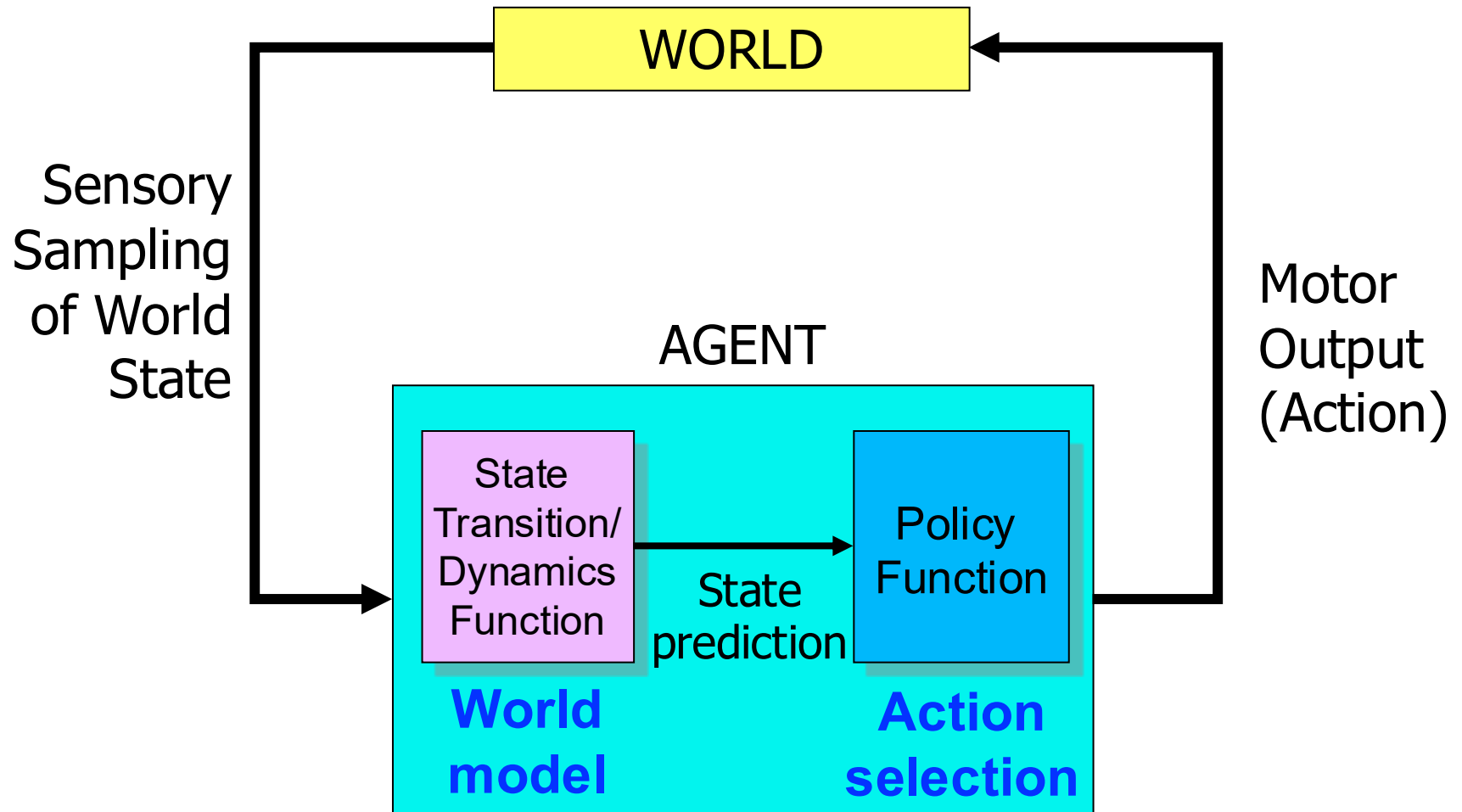
Almost all Cortical Areas have Motor Outputs



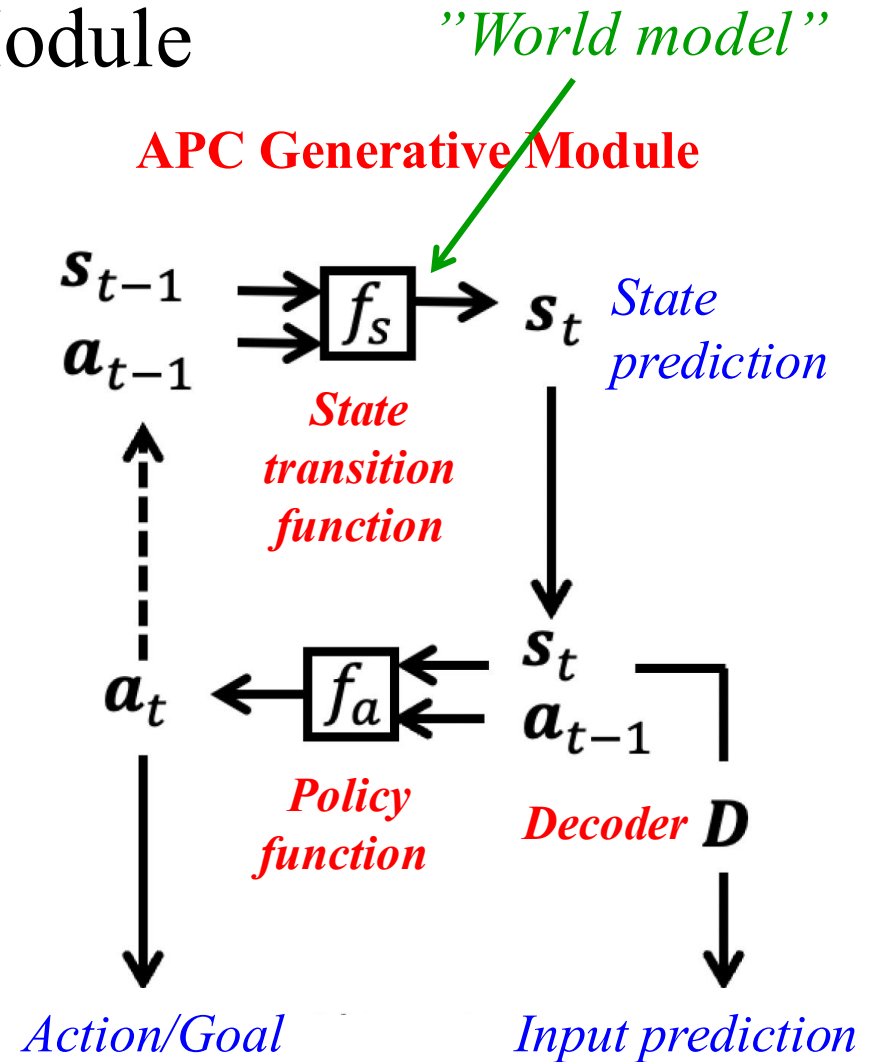
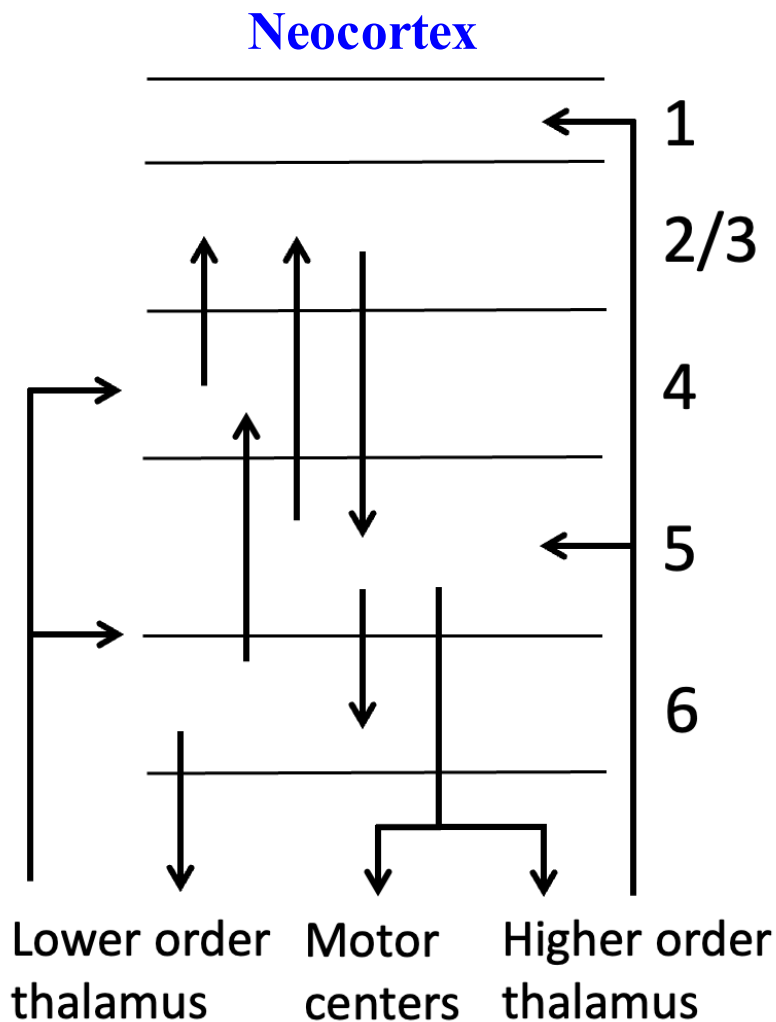
Even primary sensory cortices like V1, A1 and S1 send **outputs** to **sub-cortical motor centers** (superior colliculus, inferior colliculus, basal ganglia, spinal cord,...)

Why?

Inspiration from Reinforcement Learning and Control Theory



Active Predictive Coding (APC): Basic Sensory-Motor Module

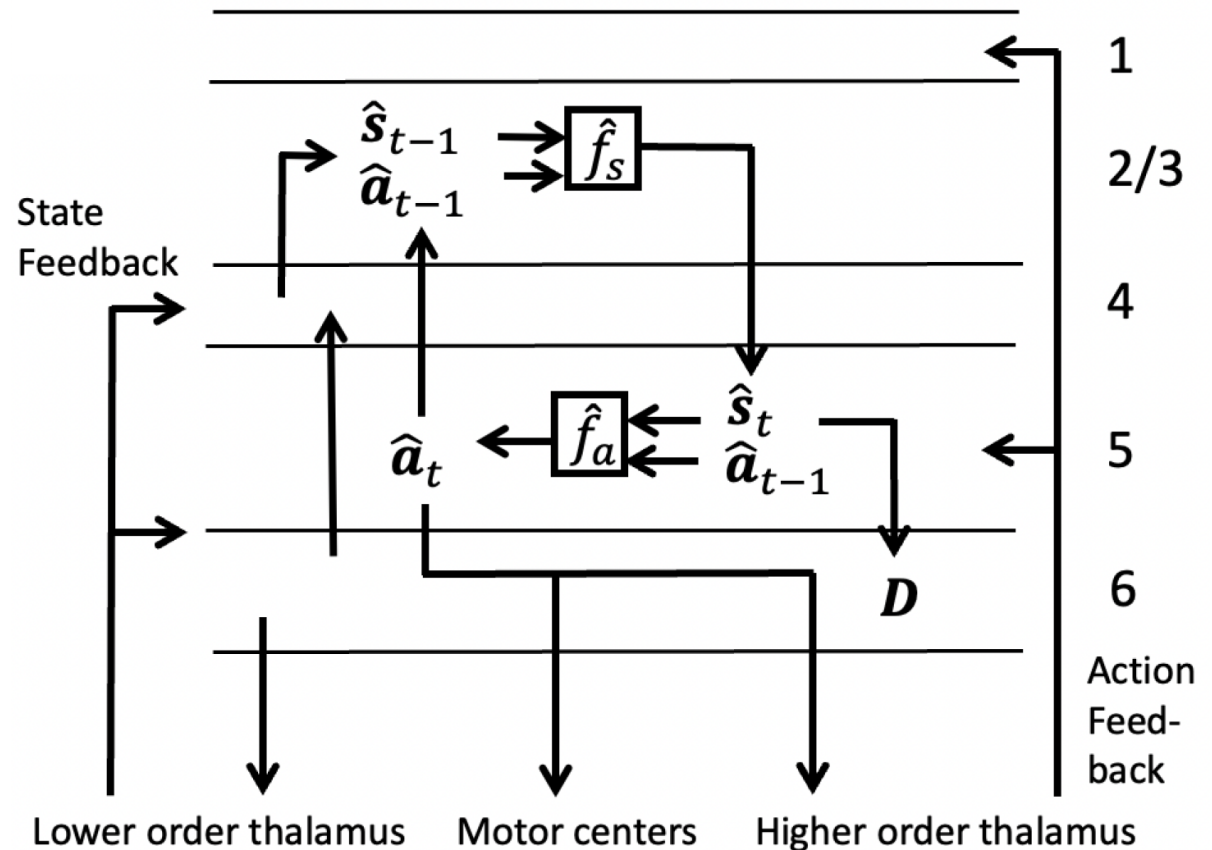


(Rao, *Nature Neuroscience*, 2024)

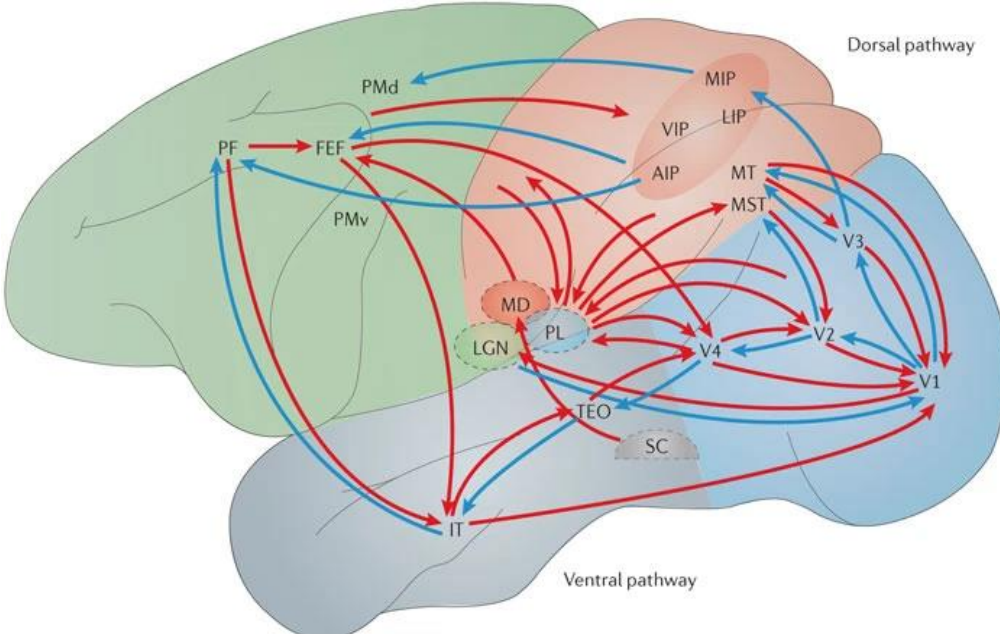
Inference and Learning in the APC Model

Inference: Each area uses feedback (e.g., prediction errors) to update its state estimate and action estimate before generating next state/action prediction

Learning: Using prediction errors and rewards

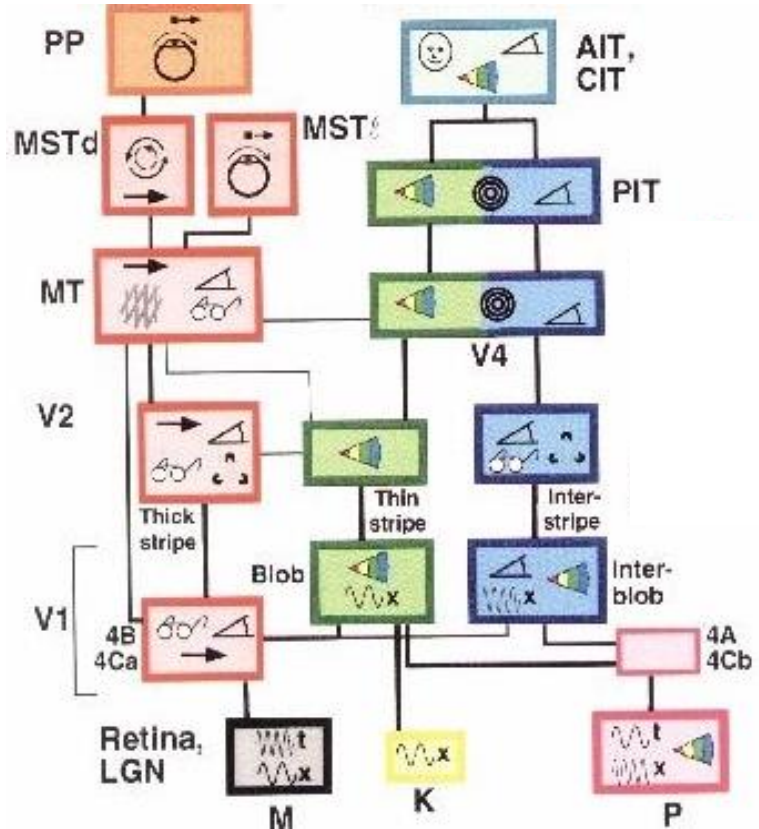


Visual hierarchy



Feedforward connections: Blue

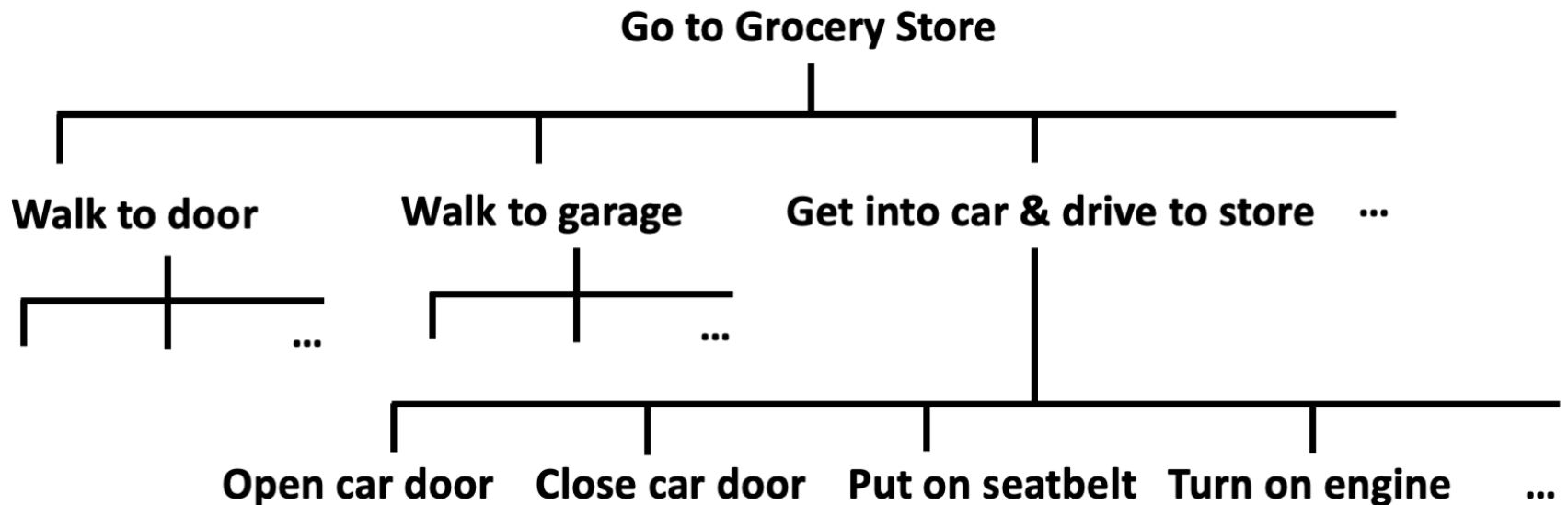
Feedback connections: Red



(Gilbert & Li, 2013; Van Essen & Maunsell, 1983; Felleman & Van Essen, 1991;...)

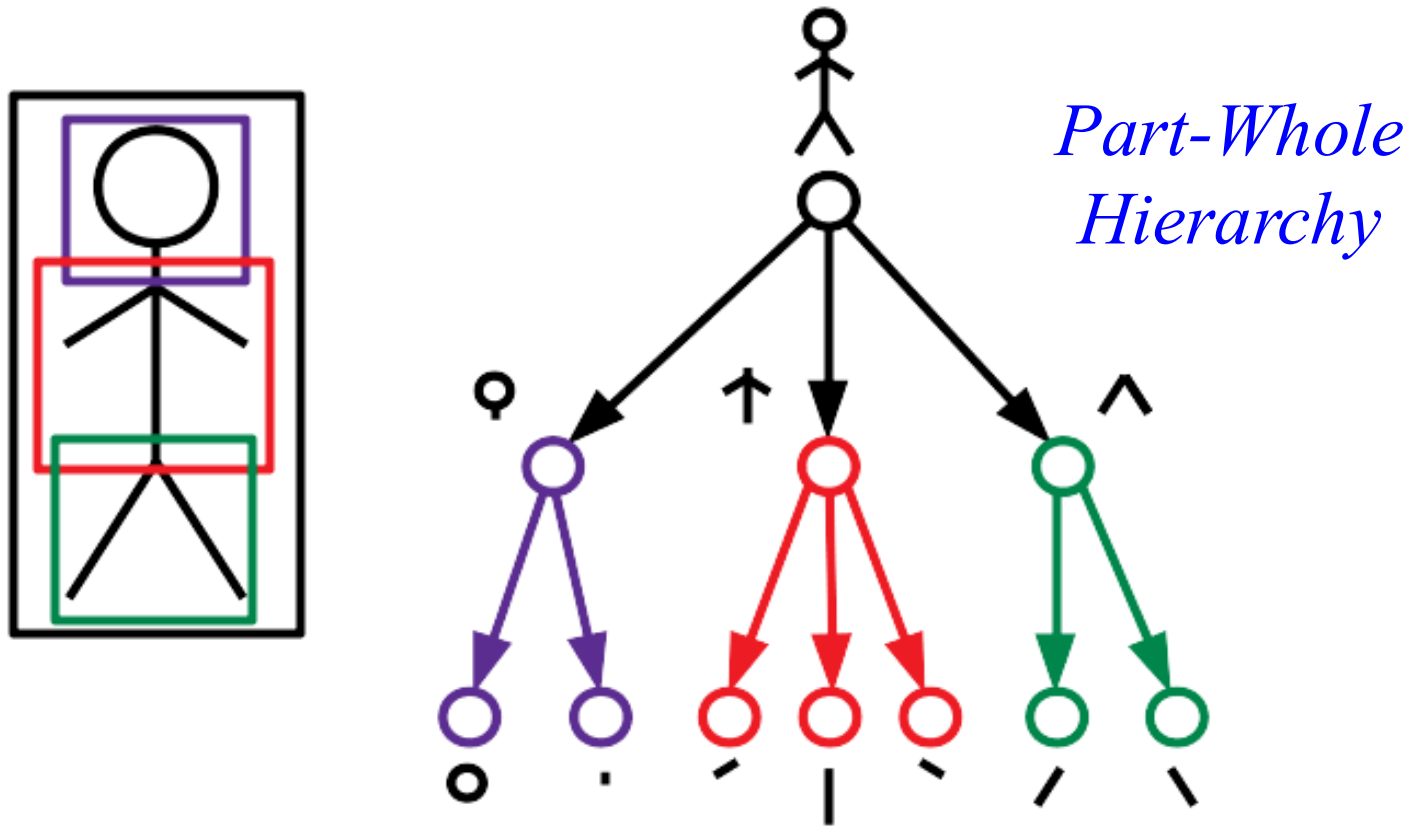
Hierarchical APC: Computational Motivation

Exploiting Compositionality to Solve Complex Problems



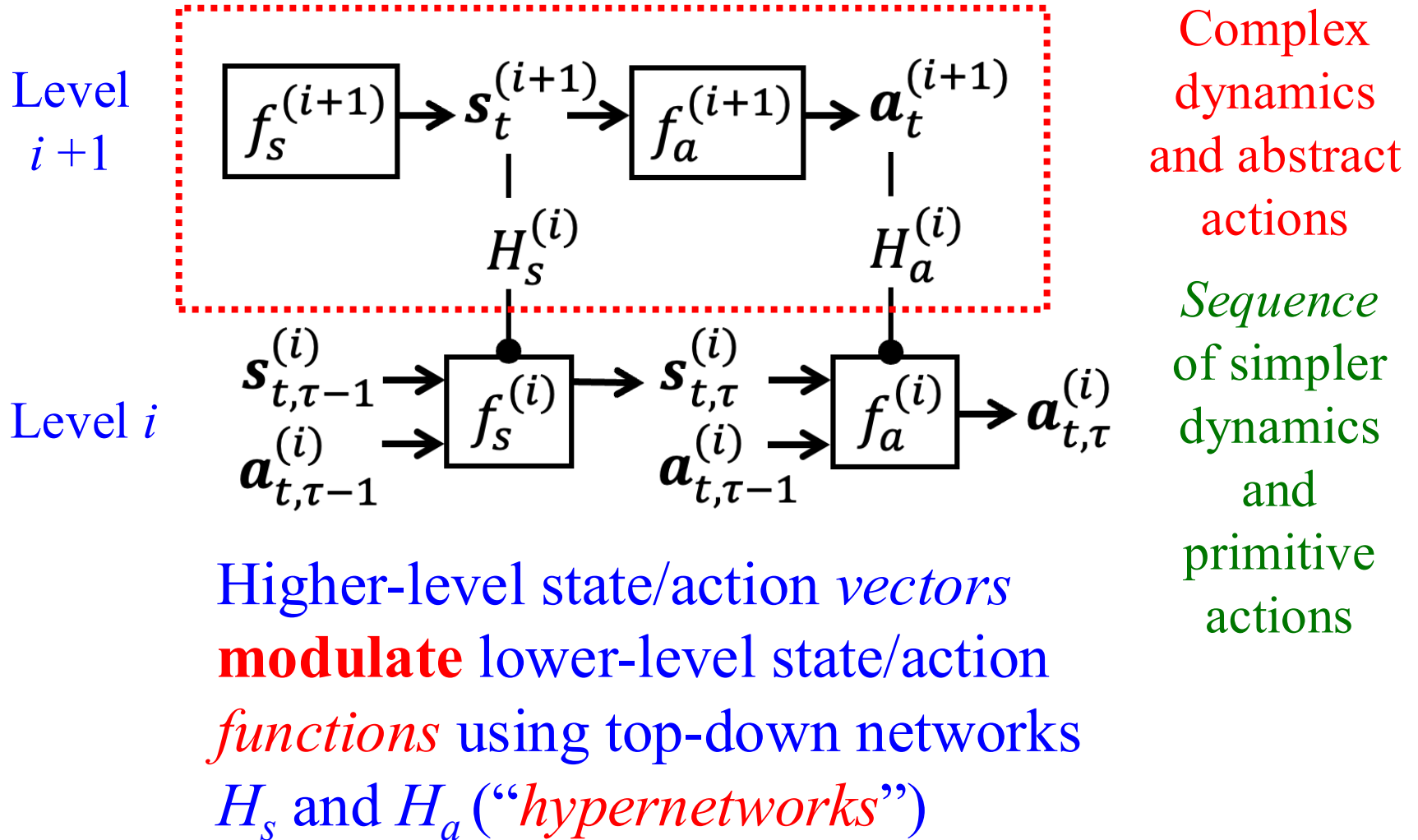
- Policies and transition functions can be re-used to solve other tasks
- Evolution builds on top of what is already there – re-uses modules

Hierarchical Representations in Perception



Parts can be re-used to represent other objects

Hierarchical APC: Implementation

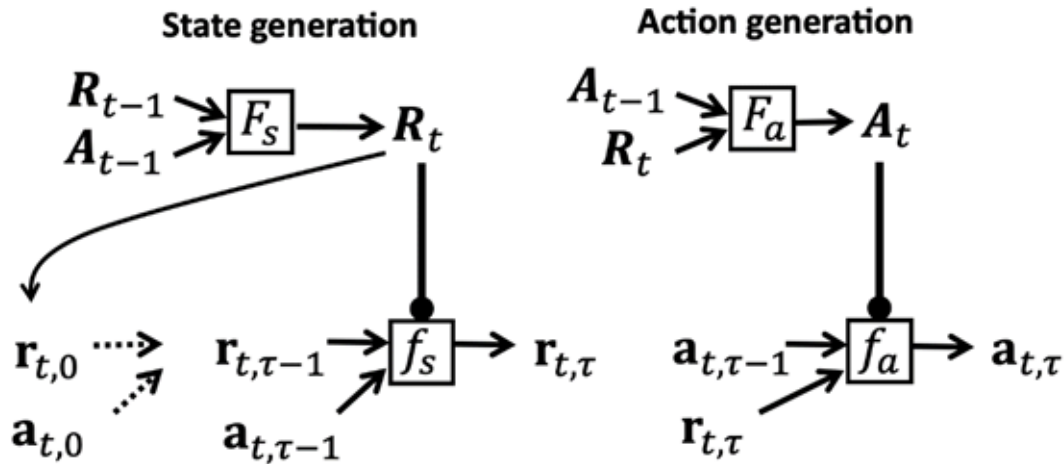


(Rao, *Nature Neuroscience*, 2024; Rao et al., *Neural Computation*, 2023)

Example #1: Active Perception

*States encode
current object/part*

*Actions encode
where to look next*

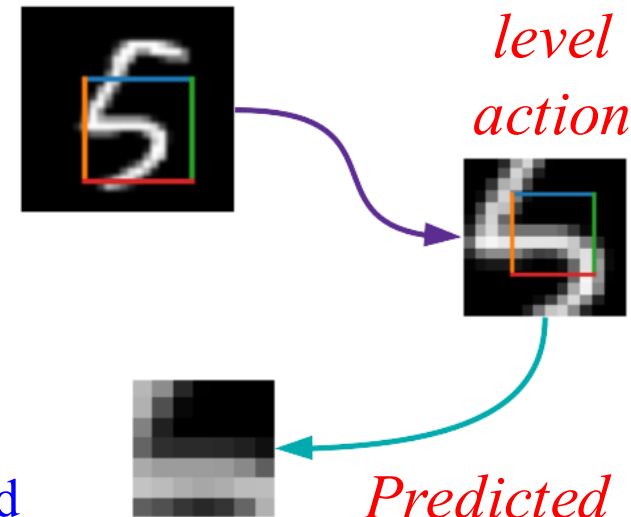


State networks trained to
minimize prediction error
via gradient descent

Action networks trained
to minimize task loss or
maximize reward via RL

*Region
selected by
higher level
action*

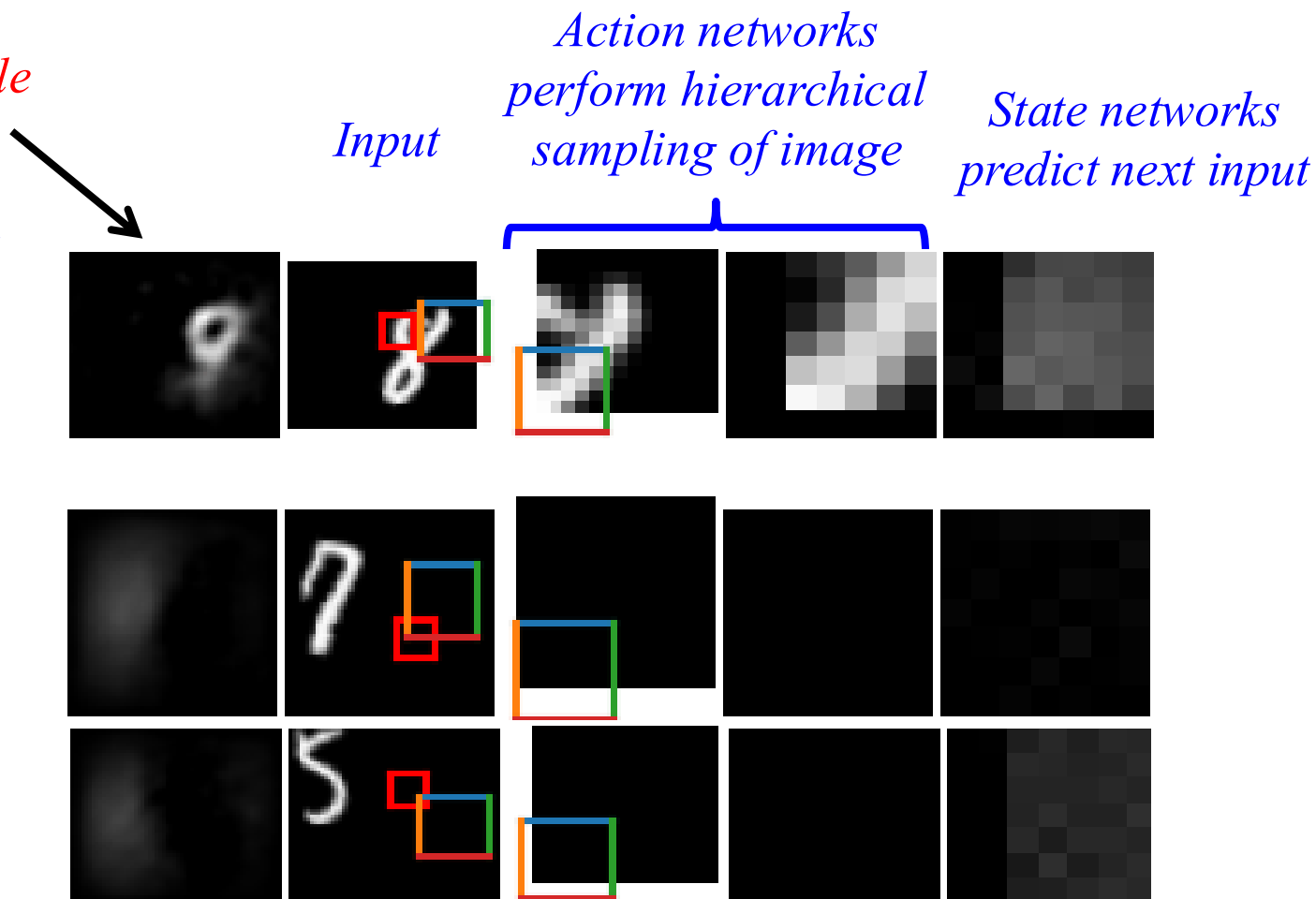
*Region
selected by lower
level
action*



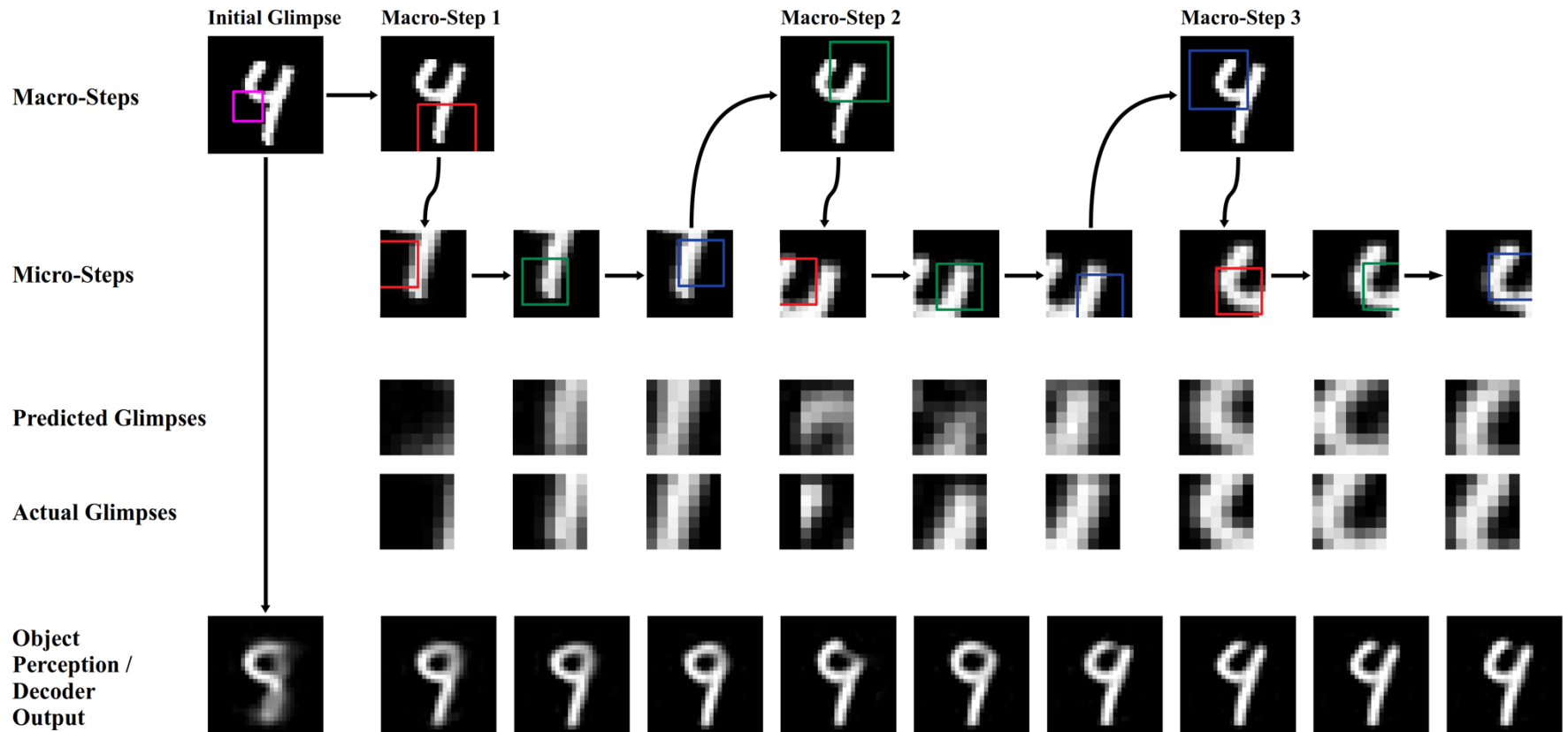
*Predicted
and analyzed
region*

Active Perception Examples (after Learning)

*Model's stable
perception
based on
accumulated
evidence*

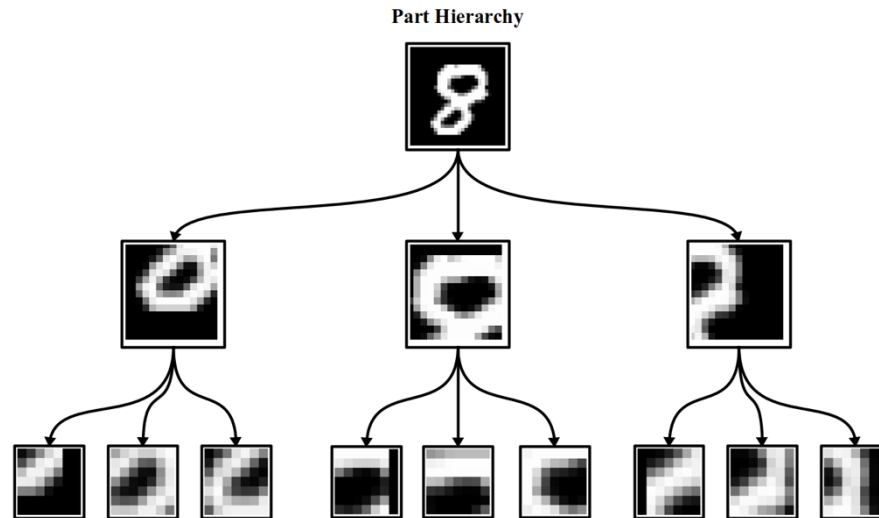


Parsing and predicting an object hierarchically

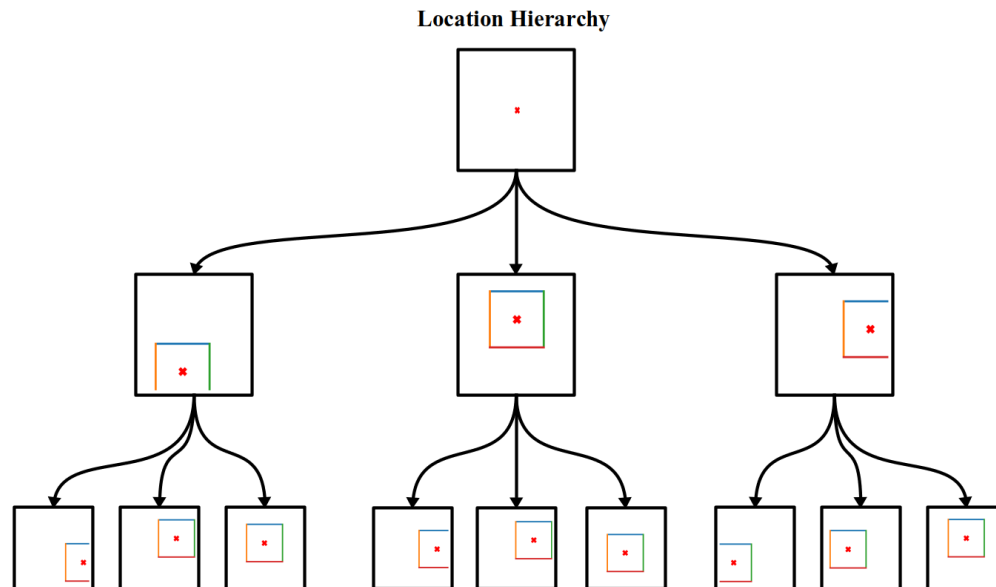


Learned Part-Whole Hierarchy for an Object

What



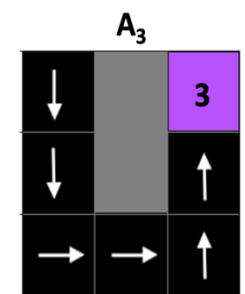
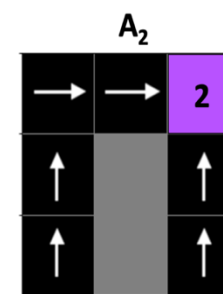
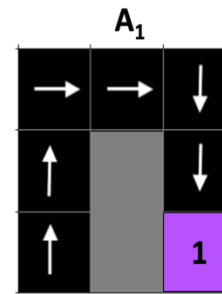
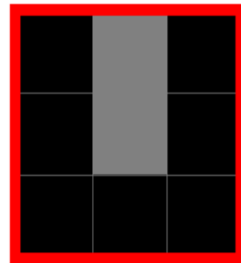
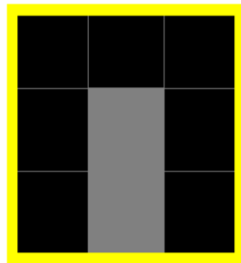
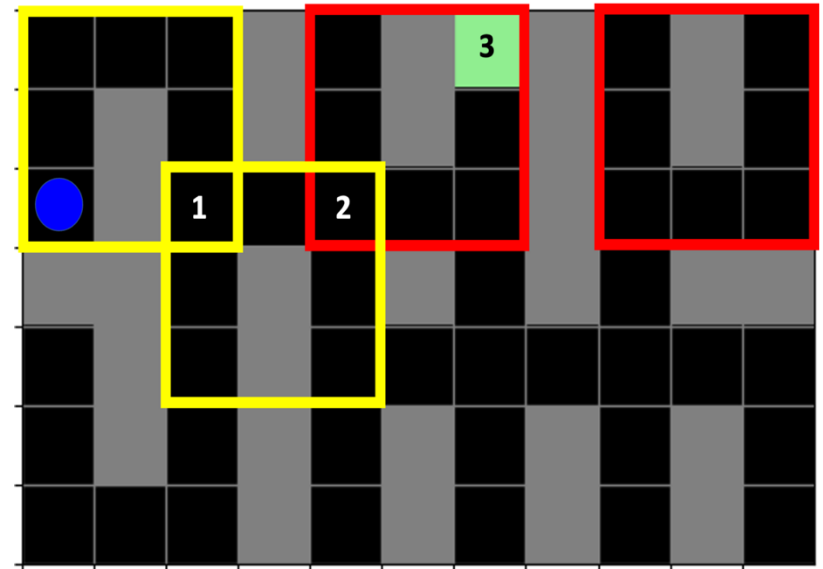
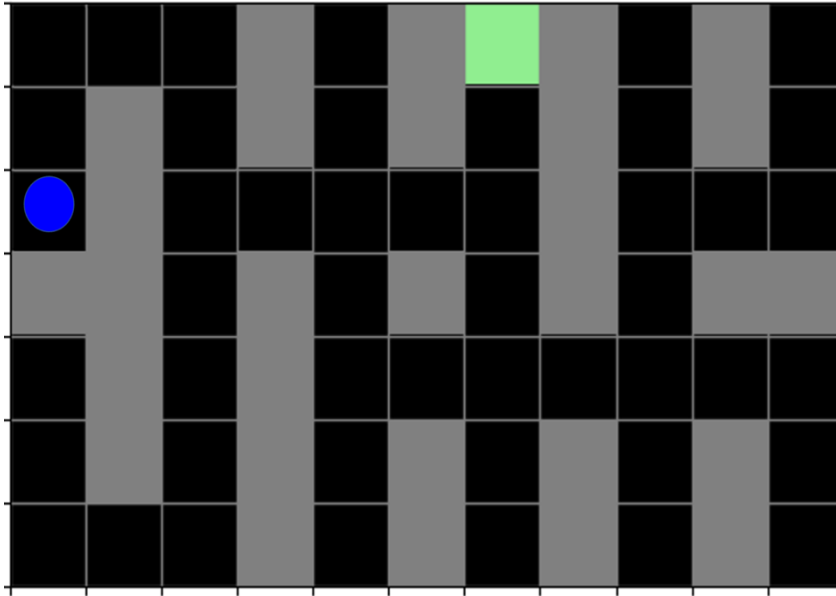
Where



Example #2: Same APC Model can be used for Hierarchical Planning!

Large “building” environment

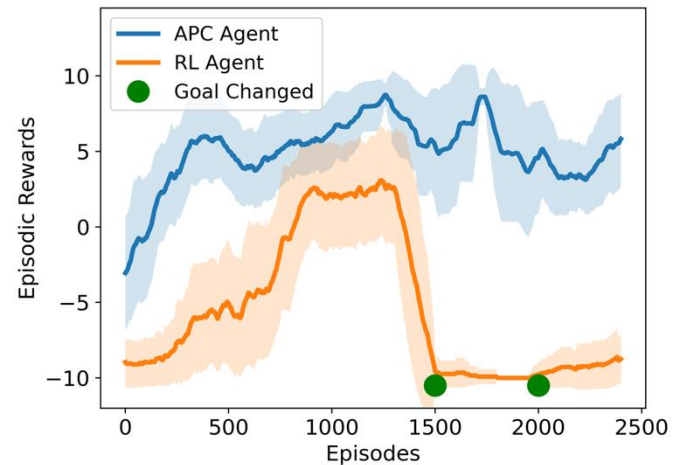
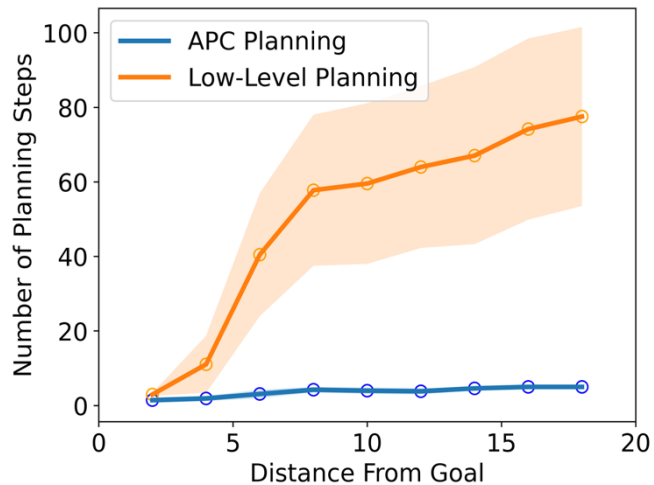
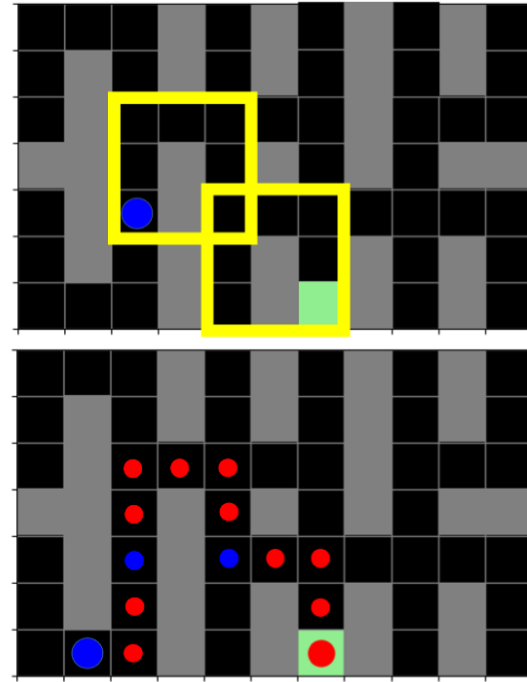
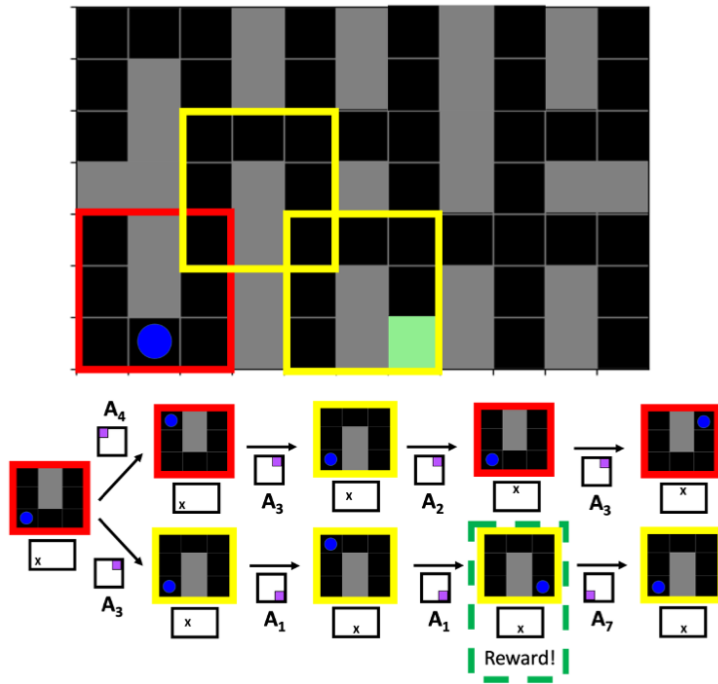
Plan using higher-level states and actions



Higher-level states = “rooms”

Higher-level actions = goals

APC Hierarchical Planning is Efficient and Adaptive

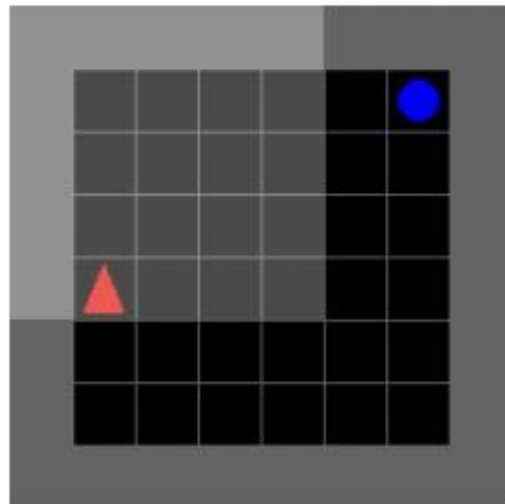


What about *Language*?

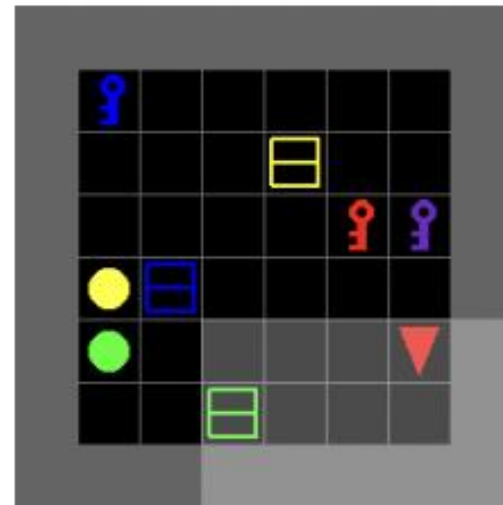
Testing the Generality of APC Framework: Language Understanding

- ❓ *Hypothesis*: If laminar cortical architecture is similar across areas/modalities, APC model used for vision and navigation should also be applicable to language understanding

BabyAI Platform: Language understanding + vision-based tasks

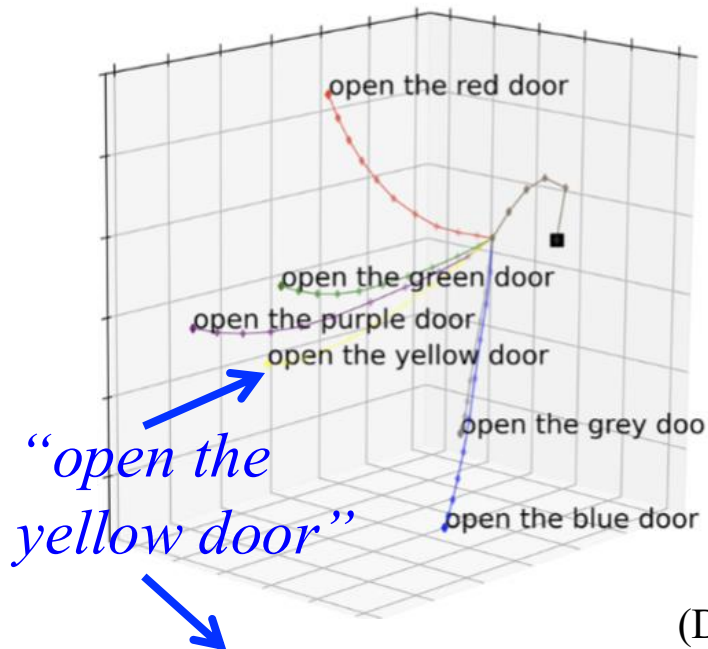


(a) GoToObj: "go to the blue ball"

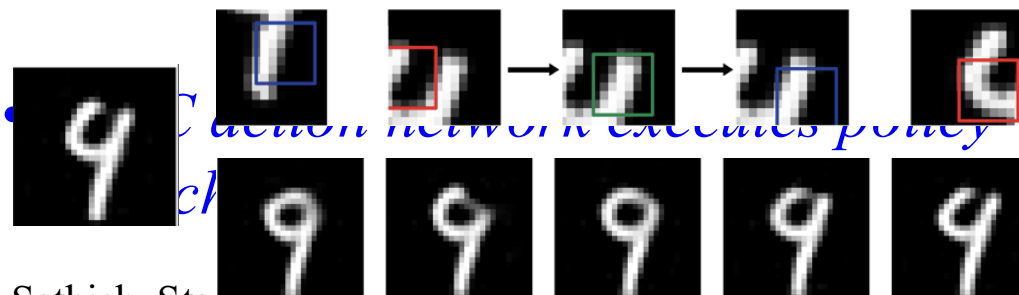


(b) PutNextLocal: "put the blue key next to the green ball"

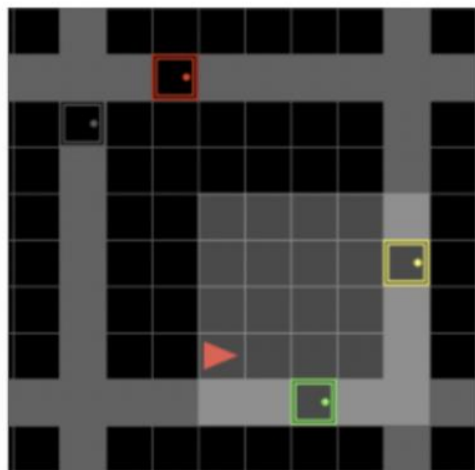
Language Understanding using APC



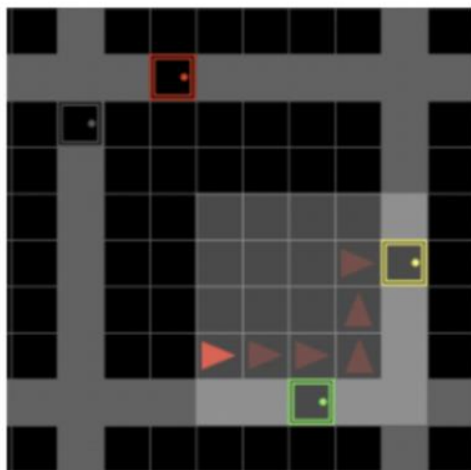
- *APC state network infers the goal of a sentence word-by-word*
 - Similar to inferring a visual object saccade-by-saccade



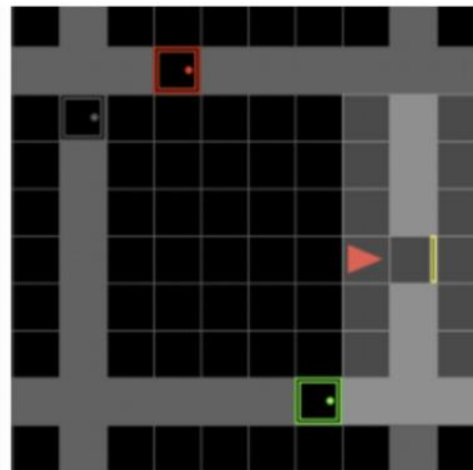
(Duan, Sathish, Stambaugh, Rao, AAAI NeuroAI Workshop, 2026)



open the yellow door



open the yellow door



open the yellow door

Summary and Conclusions

- ❑ (Active) Predictive Coding as an implementation of Bayesian inference based on an **internal generative model of the world**
 - ❑ Sensation and action **intimately intertwined** and **hierarchically organized**
- ❑ APC model suggests how the **same cortical architecture can solve a variety of problems**:
 - ❑ Active perception
 - ❑ Part-whole learning
 - ❑ Navigation & Planning
 - ❑ Language understanding
 - ❑ (Possibly) abstract concept learning
- ❑ Ongoing work: Scaling up the model via transformers and GPUs, collaborations with neuroscientists to test theory

Publications and Podcast

nature neuroscience

Perspective

<https://doi.org/10.1038/s41593-024-01673-9>

A sensory–motor theory of the neocortex

Received: 21 June 2023

Rajesh P. N. Rao  

Accepted: 26 April 2024

Neural Computation 36, 1–32 (2024)

ARTICLE Communicated by Rafal Bogacz

Active Predictive Coding: A Unifying Neural Model for Active Perception, Compositional Learning, and Hierarchical Planning

Rajesh P. N. Rao

rao@cs.washington.edu

Dimitrios C. Gklezacos

gklezd@cs.washington.edu

Vishwas Sathish

vsathish@cs.washington.edu

Lessons from Neuroscience for AI: How integrating Actions, Compositional Structure and Episodic Memory could enable Safe, Interpretable and Human-Like AI

Rajesh P.N. Rao^{1*}, Vishwas Sathish¹, Linxing Preston Jiang¹,
Matthew Bryan¹, Prashant Rangarajan¹

*Nature
Machine
Intelligence
(in review)
2026*

PLOS COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

Dynamic predictive coding: A model of hierarchical sequence learning and prediction in the neocortex

Linxing Preston Jiang ^{1,2,3}, Rajesh P. N. Rao ^{1,2,3*}



PNAS Nexus, 2023, 2, 1–9

<https://doi.org/10.1093/pnasnexus/pgad337>

Advance access publication 14 October 2023

Research Report

Recursive neural programs: A differentiable framework for learning compositional part-whole hierarchies and image grammars

Ares Fisher and Rajesh P. N. Rao 

Podcast



Acknowledgement:

“Predictive Coding Task Force”

@ Neural Systems Lab, UW



*Dimi Gklezakos
(recent PhD)*



*Preston Jiang
(recent PhD)*



*Vishwas
Sathish
(PhD student)*



*Prashant
Rangarajan
(PhD student)*



*Jeffrey Duan
(PhD student)*



*Danny
Muir
(Undergrad)*

Acknowledgments

Collaborators Past and Present

**Dana Ballard, Terry Sejnowski, Michael Shadlen, Roozbeh Kiani,
Nick Steinmetz, Jean-Claude Dreher, Krishna Jayant,
Karl Friston, Alex Ororbia, and many others**
Center for Neurotech: www.centerforneurotech.org

Funding

**NSF, NIH, ONR, DARPA, TWCF, AFOSR,
Weill Neurohub, Hwang professorship**