Universität Zürich[UZH] ETH zürich

Translational Neuromodeling Unit

Predictive Coding

Computational Psychiatry Course Zurich

14.09.2022

Alex Hess

Translational Neuromodeling Unit

University of Zurich & ETH Zurich

“Bayesian brain” hypothesis

forward model

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

x: state of the world
y: sensory input
m: current model

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

perception

Bayes’ rule

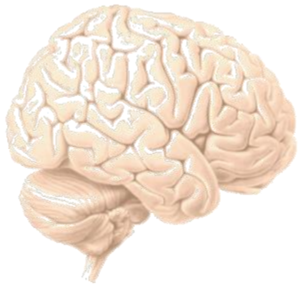
$$p(x|y,m) = \frac{p(y|x,m)p(x|m)}{p(y|m)}$$

Figure adapted from a slide by Klaas Enno Stephan

"Bayesian brain" hypothesis

forward model

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



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

perception

x: state of the world
y: sensory input
m: current model

Bayes' rule

$$\underbrace{p(x|y,m)}_{\text{posterior}} = \frac{\overbrace{p(y|x,m)}^{\text{likelihood}}\overbrace{p(x|m)}^{\text{prior}}}{\underbrace{p(y|m)}_{\text{model evidence}}}$$

Figure adapted from a slide by Klaas Enno Stephan

“Bayesian brain” hypothesis

forward model

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

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

perception

x: state of the world
y: sensory input
m: current model

Bayes’ rule

$$\overbrace{p(x|y,m)}^{\text{posterior}} = \frac{\overbrace{p(y|x,m)}^{\text{likelihood}}\overbrace{p(x|m)}^{\text{prior}}}{\underbrace{p(y|m)}_{\text{model evidence}}} = \frac{p(y|x,m)p(x|m)}{\int p(y|x,m)p(x|m)dx}$$

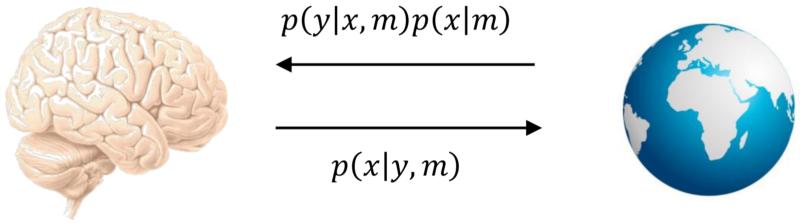
Figure adapted from a slide by Klaas Enno Stephan

(Bayesian) Predictive Coding

what?

how?

implementation?



Bayes' rule

$$p(x|y,m) = \frac{p(y|x,m)p(x|m)}{p(y|m)}$$

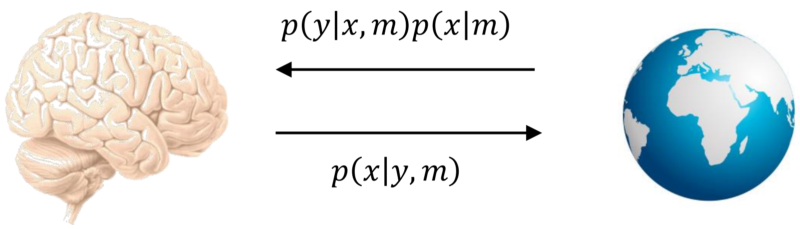
$$= \frac{p(y|x,m)p(x|m)}{\int p(y|x,m)p(x|m)dx}$$

Figure adapted from a slide by Klaas Enno Stephan

(Bayesian) Predictive Coding

Marr's levels of analysis

computational	(approximate) Bayesian inference
algorithmic	predictive coding
implementational	[predictive coding in the brain]



Bayes' rule

$$p(x|y,m) = \frac{p(y|x,m)p(x|m)}{p(y|m)}$$

$$= \frac{p(y|x,m)p(x|m)}{\int p(y|x,m)p(x|m)dx}$$

Figure adapted from a slide by Klaas Enno Stephan

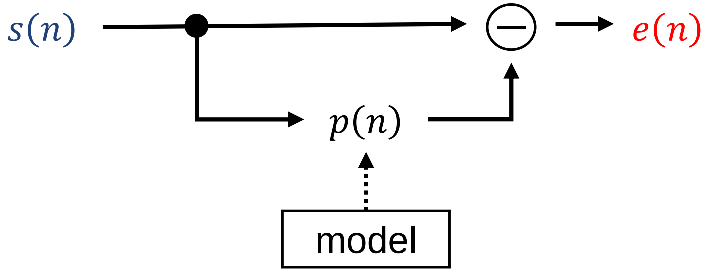
PC in engineering and information theory

$$e(n) = s(n) - p(n)$$

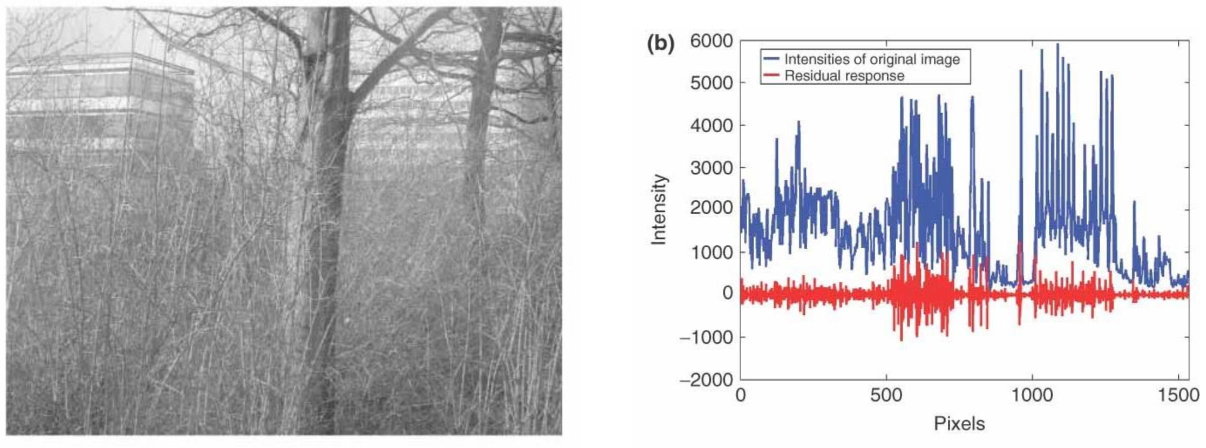
Minimum redundancy principle
(Barlow, 1961)

- Efficient way to transmit a signal $s(n)$:
  - Model $\Rightarrow$ prediction $p(n)$
  - Residual error $e(n)$

$\left.\right\}$ reconstruct signal $s(n)$



Adapted from O’Shaughnessy 1988, *IEEE Potentials*



Huang & Rao 2011, *WIREs Cogn Sci*

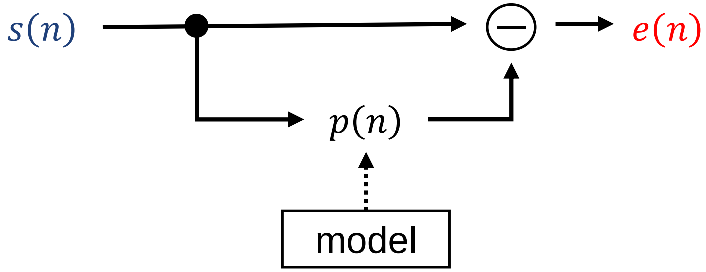
PC in engineering and information theory

$$e(n) = s(n) - p(n)$$

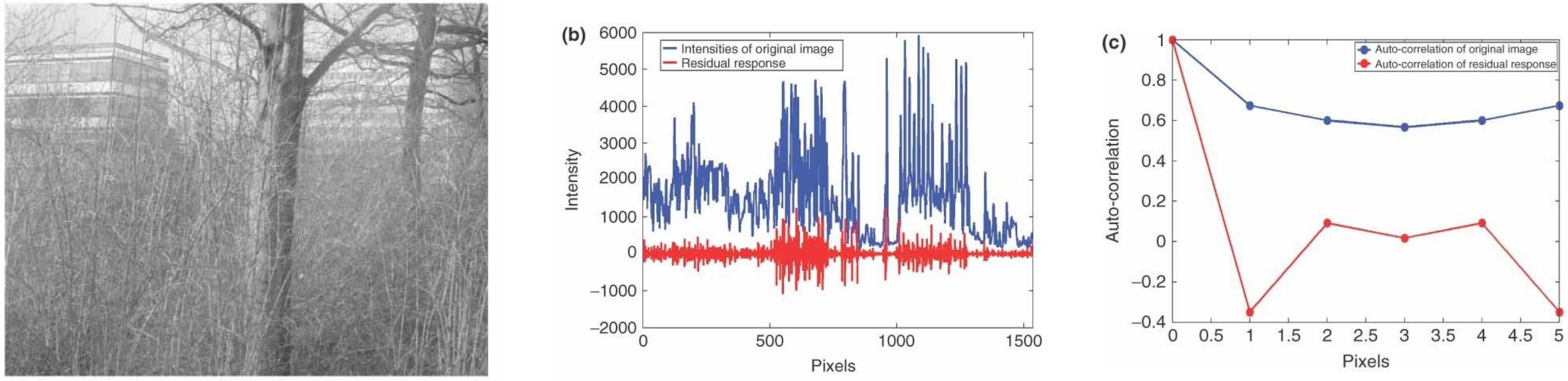
Minimum redundancy principle
(Barlow, 1961)

- Efficient way to transmit a signal $s(n)$:
  - Model $\Rightarrow$ prediction $p(n)$
  - Residual error $e(n)$

  } reconstruct signal $s(n)$
- Decorrelation

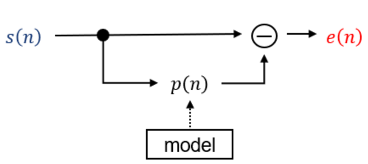


Adapted from O'Shaughnessy 1988, *IEEE Potentials*



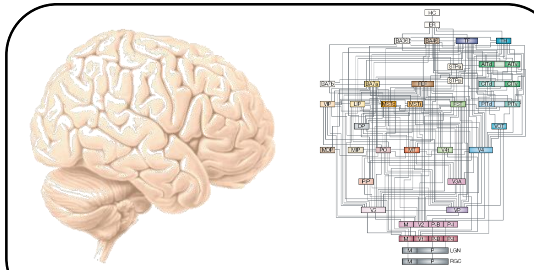
Huang & Rao 2011, *WIREs Cogn Sci*

Predictive Coding as neuroscientific theory



Intellectual antecedents

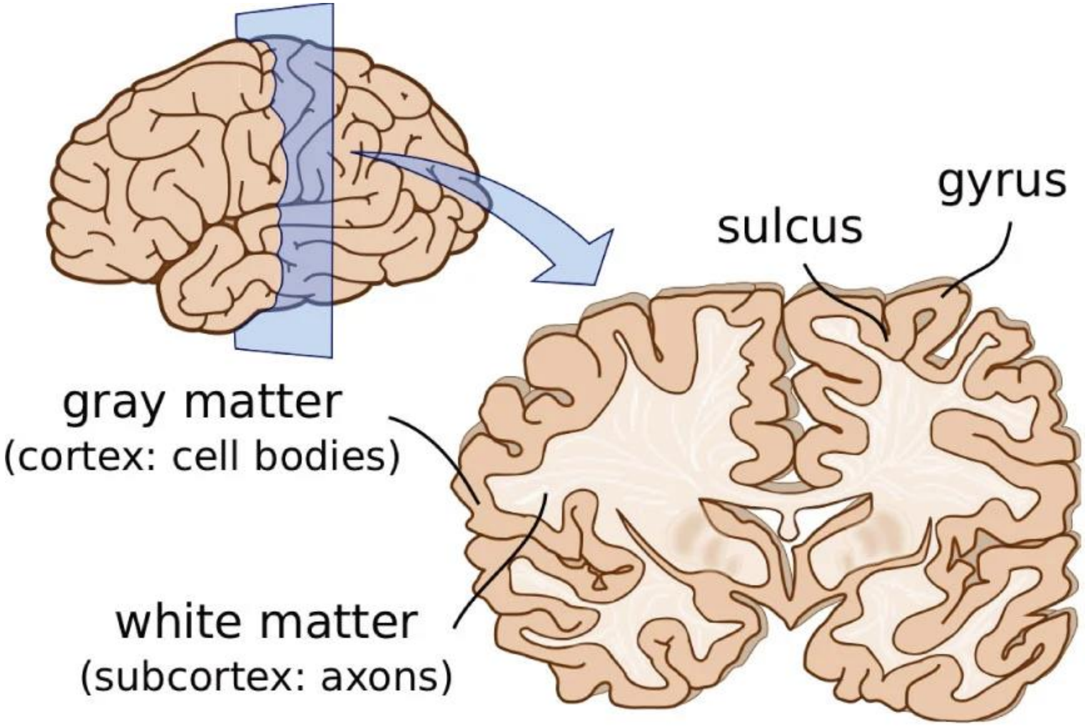
- Redundancy reduction



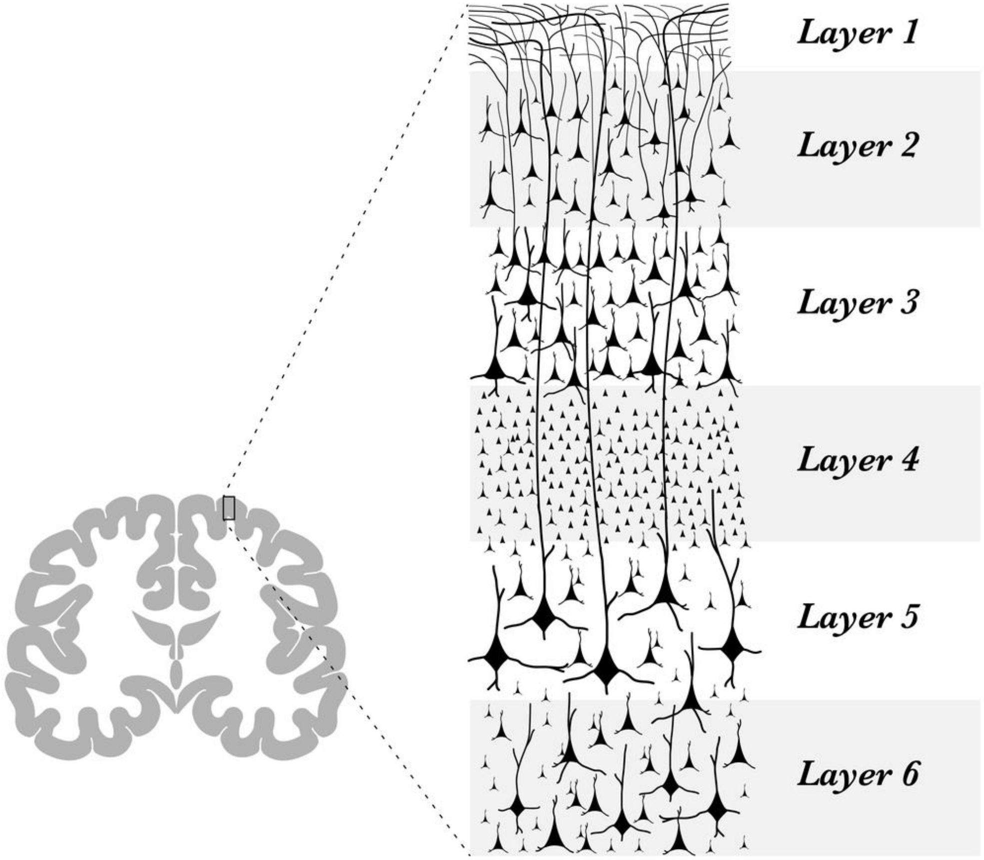
Felleman & Van Essen 1991, *Cereb Cortex*

Neuroanatomy

Cerebral Cortex



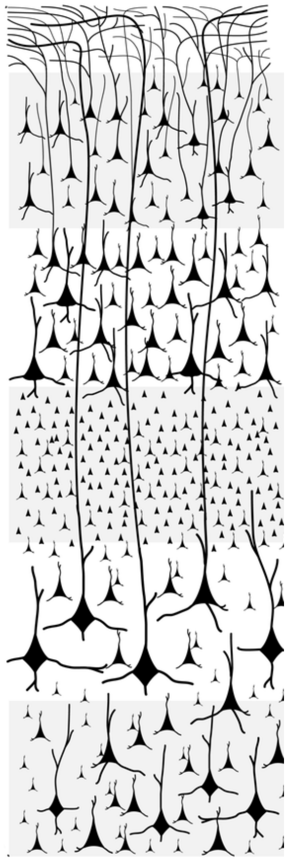
Budday et al. 2014, *Sci Rep*



Barrett 2017

Cell layers of the neocortex

	Lamina …	
Supragranular layers	*Molecularis*	I
	Granularis Externa	II
	Pyramidialis Externa	III
Granular layer	*Granularis Interna*	IV
Infragranular layers	*Pyramidialis Interna*	V
	Multiformis	VI

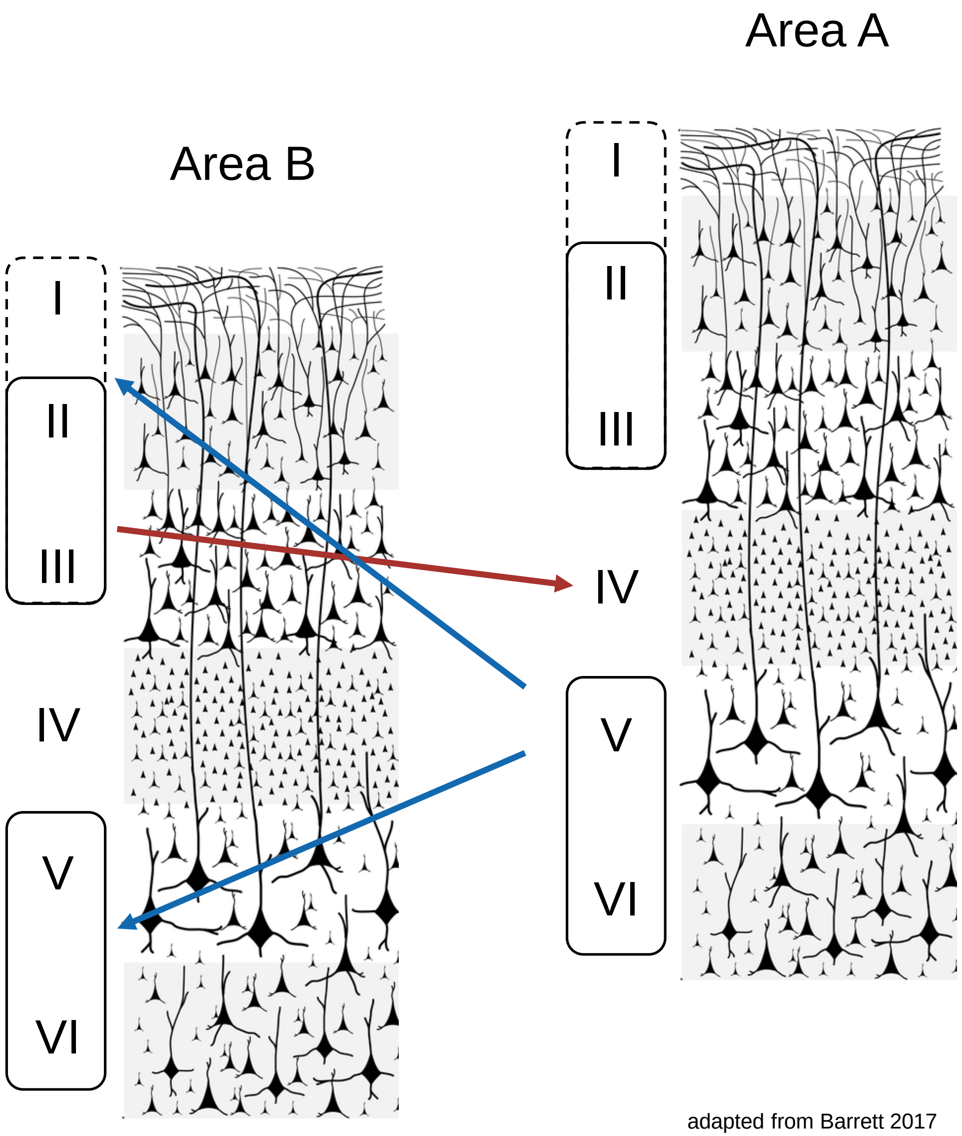


adapted from Barrett 2017

Hierarchical Relationships in the Visual Cortex

Visual cortex of macaque monkeys
(Felleman & Van Essen 1991, *Cereb Cortex*)

- Reciprocity of cortico-cortical connections
- Laminar patterns
  - Forward connections (ascending pathways):
    - Origin: superficial pyramidal cells (layers II & III)
    - Termination: granular layer (IV)
  - Backward connections (descending pathways):
    - Origin: deep pyramidal cells (layer V)
    - Termination: agranular layers (mainly I & VI)

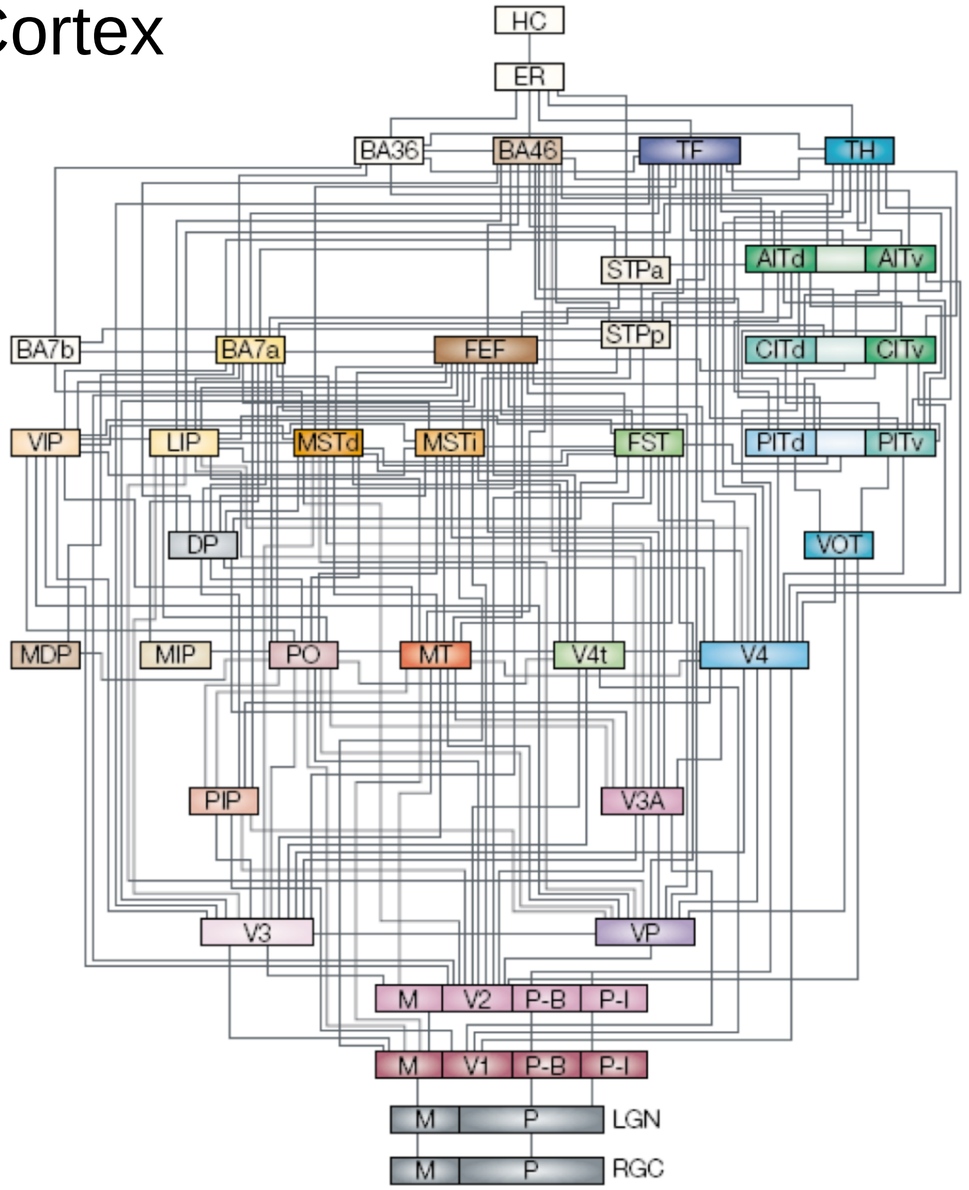


adapted from Barrett 2017

Hierarchical Relationships in the Visual Cortex

Visual cortex of macaque monkeys
(Felleman & Van Essen 1991, *Cereb Cortex*)

- Reciprocity of cortico-cortical connections
- Laminar patterns
  - Forward connections (ascending pathways):
    - Origin: superficial pyramidal cells (layers II & III)
    - Termination: granular layer (IV)
  - Backward connections (descending pathways):
    - Origin: deep pyramidal cells (layer V)
    - Termination: agranular layers (mainly I & VI)
- Identify hierarchy based on laminar patterns of cortical connectivity (forward & backward connections)
- Hierarchical relationships also…
  - In other regions (somatosensory, auditory cortex, etc.)
  - In other species (other primates, cats, rats, etc.)

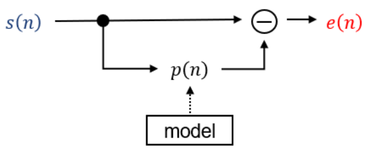


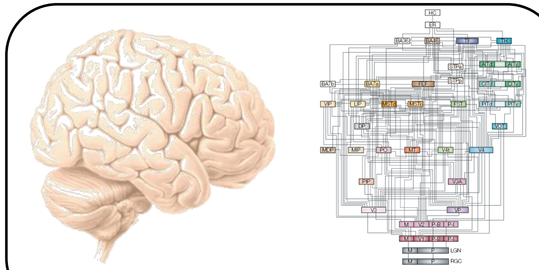
Felleman & Van Essen 1991, *Cereb Cortex*

Predictive Coding as neuroscientific theory

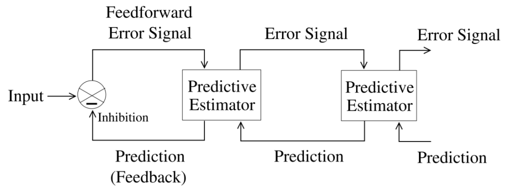
On the computational architecture of the neocortex
(D. Mumford 1992, *Biol Cybern*)

Predictive coding in the visual cortex
(Rao & Ballard 1999, *Nat Neurosci*)





Felleman & Van Essen 1991, *Cereb Cortex*



(Rao & Ballard 1999, *Nat Neurosci*)

Intellectual antecedents

- Redundancy reduction

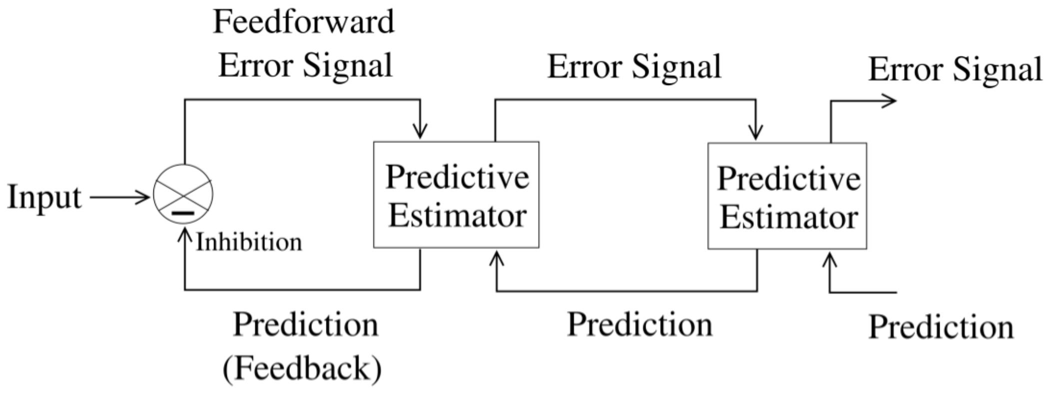
Neuroanatomy

- Hierarchical organization of cortex

Hierarchical PC model

Hierarchical predictive coding model

- Hierarchical network
  - Feedback connections: predictions
  - Feedforward connections: error signal
  - Predictive estimator: use error signal to generate next prediction



$$\boldsymbol{I} = f(U\boldsymbol{r}) + \boldsymbol{n}$$

$$\mathbf{r} = \boldsymbol{r}^{td} + \boldsymbol{n}^{td}$$
$$= f\left(U^h \boldsymbol{r}^h\right) + \boldsymbol{n}^{td}$$

$\boldsymbol{I}$: inputs

$\boldsymbol{r}$: causes

U: weights

f: activation function

$\boldsymbol{n}$: noise

U^h: higher-level weights

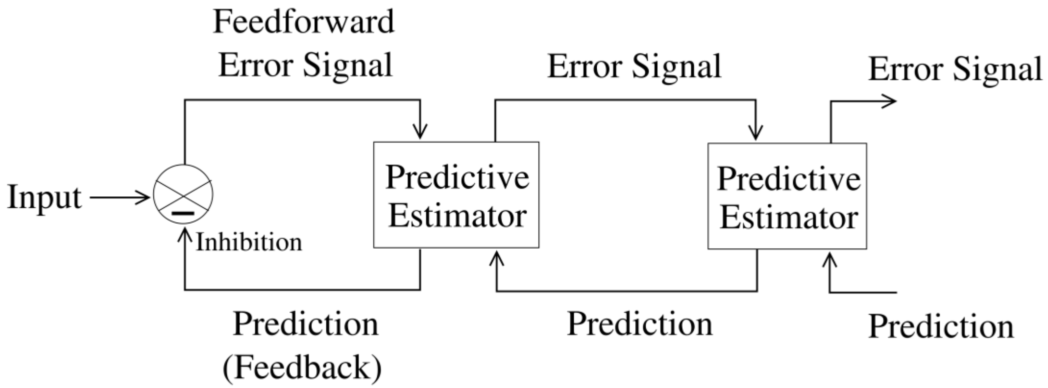
$\boldsymbol{r}^h$: higher-level causes

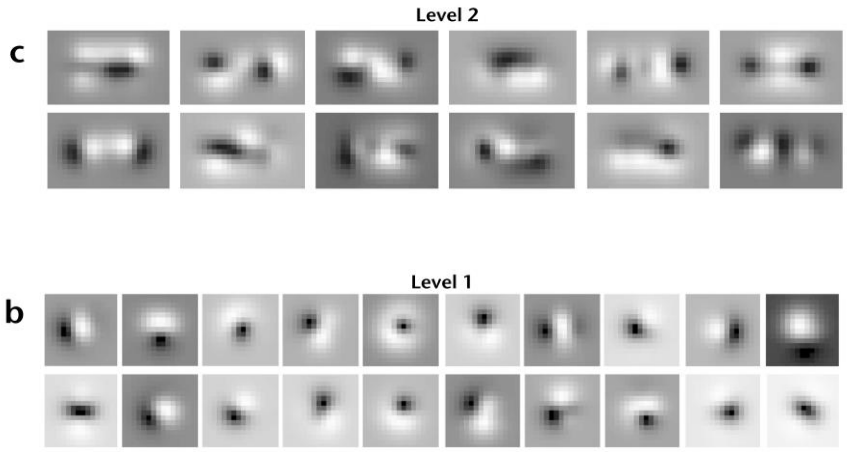
$\boldsymbol{n}^{td}$: noise

Rao & Ballard 1999, *Nat Neurosci*

Hierarchical predictive coding model

- Hierarchical network
  - Feedback connections: predictions
  - Feedforward connections: error signal
  - Predictive estimator: use error signal to generate next prediction
- Train network on patches of static natural images
  - Learned synaptic weights resemble cell-like receptive fields
  - Receptive field sizes: lower vs. upper levels

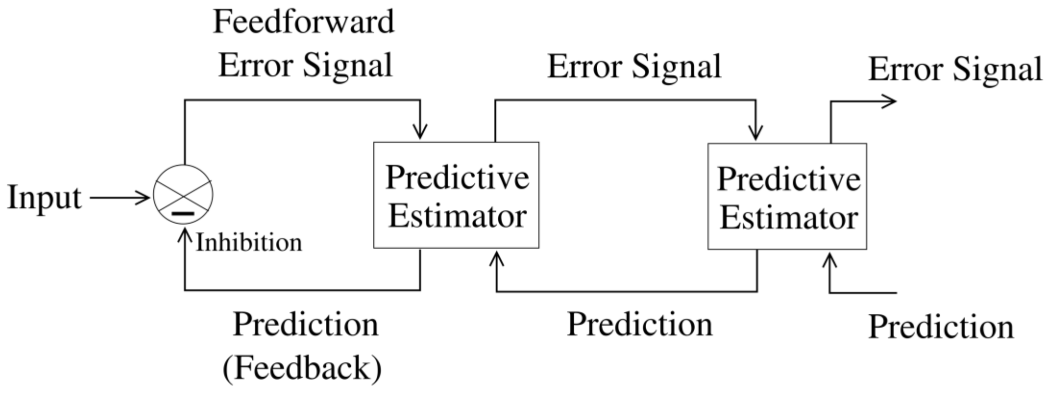


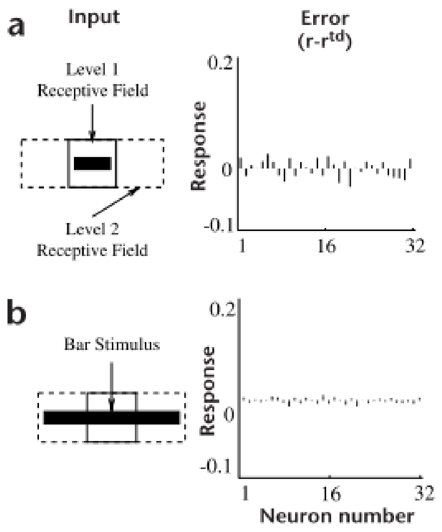


Rao & Ballard 1999, *Nat Neurosci*

Hierarchical predictive coding model

- Hierarchical network
  - Feedback connections: predictions
  - Feedforward connections: error signal
  - Predictive estimator: use error signal to generate next prediction
- Train network on patches of static natural images
  - Learned synaptic weights resemble cell-like receptive fields
  - Receptive field sizes: lower vs. upper levels
- Functional explanation for extra-classical receptive field effects:
  - Endstopping: error-detecting model neurons





Rao & Ballard 1999, *Nat Neurosci*

Hierarchical predictive coding model

- Assume probabilistic hierarchical generative model for images
  - Cost function: negative log joint (⇒ MAP estimation)

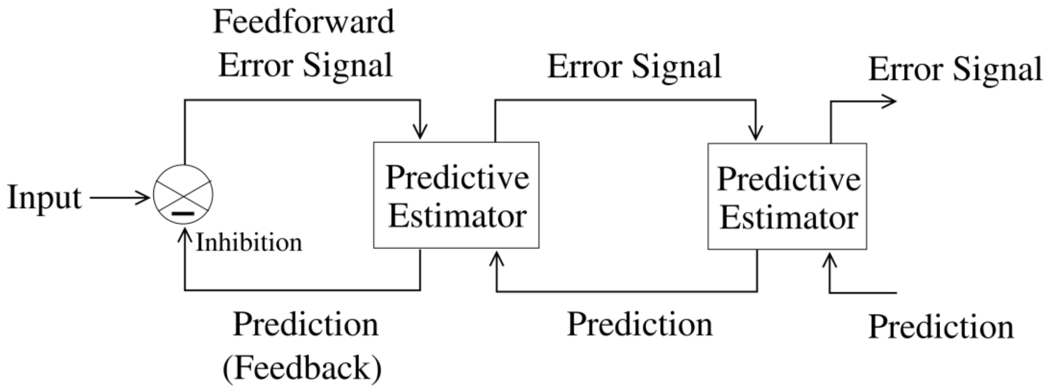
$$\frac{1}{\sigma^2}(\boldsymbol{I} - f(U\boldsymbol{r}))^T(\boldsymbol{I} - f(U\boldsymbol{r})) + \frac{1}{\sigma_{td}^2}(\boldsymbol{r} - \boldsymbol{r}^{td})^T(\boldsymbol{r} - \boldsymbol{r}^{td})$$

$$E = -\log p(\boldsymbol{I}|\boldsymbol{r}, U) - \log p(\boldsymbol{r}) - \log p(U)$$

$$= -\log(p(\boldsymbol{I}|\boldsymbol{r}, U)\, p(\boldsymbol{r})\, p(U))$$

posterior ∝ likelihood * prior

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



$$\boldsymbol{I} = f(U\boldsymbol{r}) + \boldsymbol{n} \qquad \mathbf{r} = \boldsymbol{r}^{td} + \boldsymbol{n}^{td}$$

$$\Rightarrow p(\boldsymbol{I}|\boldsymbol{r}, U)$$

$\boldsymbol{I}$: inputs

$\boldsymbol{r}$: causes

U: weights

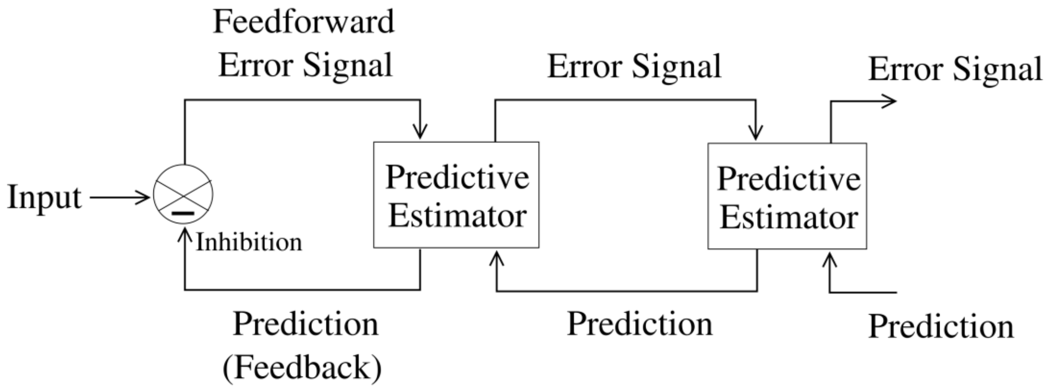
f: activation function

$\boldsymbol{n}$: noise

Rao & Ballard 1999, *Nat Neurosci*

Hierarchical predictive coding model

- Assume probabilistic hierarchical generative model for images
  - Cost function: negative log joint (⇒ MAP estimation)
- Network dynamics & synaptic learning rules
  - Error signal weighted by inverse variances (precisions)



$$\boldsymbol{I} = f(U\boldsymbol{r}) + \boldsymbol{n} \qquad \mathbf{r} = \boldsymbol{r}^{td} + \boldsymbol{n}^{td}$$

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = -\frac{k_1}{2}\frac{\partial E}{\partial \boldsymbol{r}}$$

$$= \frac{k_1}{\sigma^2} U^T \frac{\partial f}{\partial U\boldsymbol{r}}^T \big(\boldsymbol{I} - f(U\boldsymbol{r})\big) + \frac{k_1}{\sigma_{td}^2}(\boldsymbol{r}^{td} - \boldsymbol{r}) - k_1 \alpha \boldsymbol{r}$$

$$\frac{\mathrm{dU}}{\mathrm{d}t} = -\frac{k_2}{2}\frac{\partial E}{\partial U} = \frac{k_2}{\sigma^2}\frac{\partial f}{\partial U\boldsymbol{r}}^T \big(\boldsymbol{I} - f(U\boldsymbol{r})\big)\boldsymbol{r}^T - \frac{k_2}{2}\lambda U$$

$\boldsymbol{I}$: inputs

$\boldsymbol{r}$: causes

U: weights

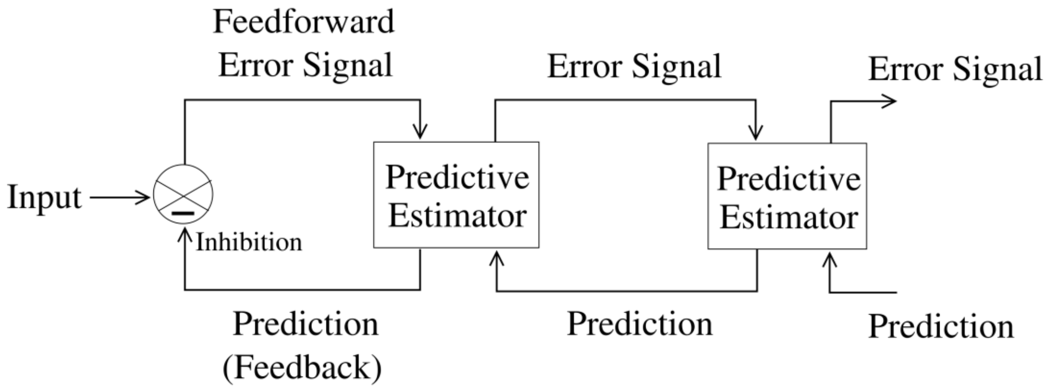
f: activation function

$\boldsymbol{n}$: noise

Rao & Ballard 1999, *Nat Neurosci*

Hierarchical predictive coding model

- Assume probabilistic hierarchical generative model for images
  - Cost function: negative log joint (⇒ MAP estimation)
- Network dynamics & synaptic learning rules
  - Error signal weighted by inverse variances (precisions)



$$\boldsymbol{I} = f(U\boldsymbol{r}) + \boldsymbol{n} \qquad \mathbf{r} = \boldsymbol{r}^{td} + \boldsymbol{n}^{td}$$

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = -\frac{k_1}{2}\frac{\partial E}{\partial \boldsymbol{r}}$$

$$= \frac{k_1}{\sigma^2} U^T \frac{\partial f}{\partial U\boldsymbol{r}}^T \left(\boldsymbol{I} - f(U\boldsymbol{r})\right) + \frac{k_1}{\sigma_{td}^2}(\boldsymbol{r}^{td} - \boldsymbol{r}) - k_1 \alpha \boldsymbol{r}$$

$$\frac{\mathrm{dU}}{\mathrm{d}t} = -\frac{k_2}{2}\frac{\partial E}{\partial U} = \frac{k_2}{\sigma^2}\frac{\partial f}{\partial U\boldsymbol{r}}^T \left(\boldsymbol{I} - f(U\boldsymbol{r})\right)\boldsymbol{r}^T - \frac{k_2}{2}\lambda U$$

$\boldsymbol{I}$: inputs

$\boldsymbol{r}$: causes

U: weights

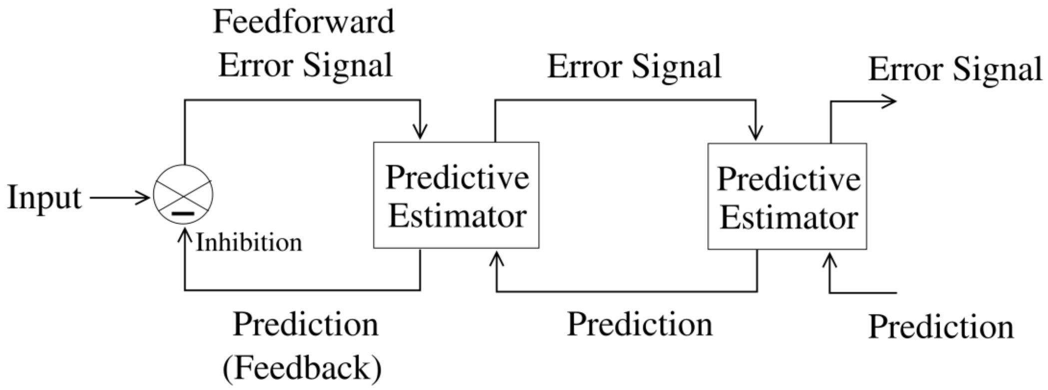
f: activation function

$\boldsymbol{n}$: noise

Rao & Ballard 1999, *Nat Neurosci*

Hierarchical predictive coding model

- Assume probabilistic hierarchical generative model for images
  - Cost function: negative log joint (⇒ MAP estimation)
- Network dynamics & synaptic learning rules
  - Error signal weighted by inverse variances (precisions)
  - Single cost function accounts for inference (updating $\boldsymbol{r}$) & learning (updating U)



$$\boldsymbol{I} = f(U\boldsymbol{r}) + \boldsymbol{n} \qquad \mathbf{r} = \boldsymbol{r}^{td} + \boldsymbol{n}^{td}$$

Δbelief ~ precision × PE

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = -\frac{k_1}{2}\frac{\partial E}{\partial \boldsymbol{r}}$$

$$= \frac{k_1}{\sigma^2} U^T \frac{\partial f}{\partial U\boldsymbol{r}}^T \big(\boldsymbol{I} - f(U\boldsymbol{r})\big) + \frac{k_1}{\sigma_{td}^2}(\boldsymbol{r}^{td} - \boldsymbol{r}) - k_1 \alpha \boldsymbol{r}$$

$$\frac{\mathrm{dU}}{\mathrm{d}t} = -\frac{k_2}{2}\frac{\partial E}{\partial U} = \frac{k_2}{\sigma^2}\frac{\partial f}{\partial U\boldsymbol{r}}^T \big(\boldsymbol{I} - f(U\boldsymbol{r})\big)\boldsymbol{r}^T - \frac{k_2}{2}\lambda U$$

$\boldsymbol{I}$: inputs

$\boldsymbol{r}$: causes

U: weights

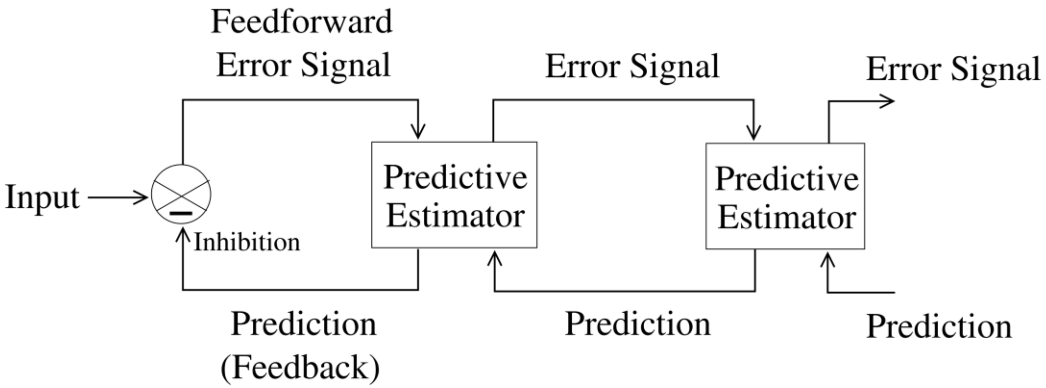
f: activation function

$\boldsymbol{n}$: noise

Rao & Ballard 1999, *Nat Neurosci*

Hierarchical predictive coding model

- Assume probabilistic hierarchical generative model for images
  - Cost function: negative log joint (⇒ MAP estimation)
- Network dynamics & synaptic learning rules
  - Error signal weighted by inverse variances (precisions)
  - Single cost function accounts for inference (updating $\boldsymbol{r}$) & learning (updating U)
  - Separation of timescales



$$\boldsymbol{I} = f(U\boldsymbol{r}) + \boldsymbol{n} \qquad \mathbf{r} = \boldsymbol{r}^{td} + \boldsymbol{n}^{td}$$

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = -\frac{k_1}{2}\frac{\partial E}{\partial \boldsymbol{r}}$$

$$= \frac{k_1}{\sigma^2} U^T \frac{\partial f}{\partial U\boldsymbol{r}}^T \big(\boldsymbol{I} - f(U\boldsymbol{r})\big) + \frac{k_1}{\sigma_{td}^2}(\boldsymbol{r}^{td} - \boldsymbol{r}) - k_1 \alpha \boldsymbol{r}$$

$$\frac{\mathrm{dU}}{\mathrm{d}t} = -\frac{k_2}{2}\frac{\partial E}{\partial U} = \frac{k_2}{\sigma^2}\frac{\partial f}{\partial U\boldsymbol{r}}^T \big(\boldsymbol{I} - f(U\boldsymbol{r})\big)\boldsymbol{r}^T - \frac{k_2}{2}\lambda U$$

$\boldsymbol{I}$: inputs

$\boldsymbol{r}$: causes

U: weights

f: activation function

$\boldsymbol{n}$: noise

Rao & Ballard 1999, *Nat Neurosci*

Predictive Coding as neuroscientific theory

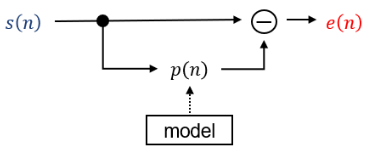
Learning and Inference in the Brain
(Friston 2003, *Neural Netw*)

A theory of cortical responses
(Friston 2005, *Phil Trans Royal Soc B*)

Hierarchical Models in the Brain
(Friston 2008, *PLoS Comput Biol*)

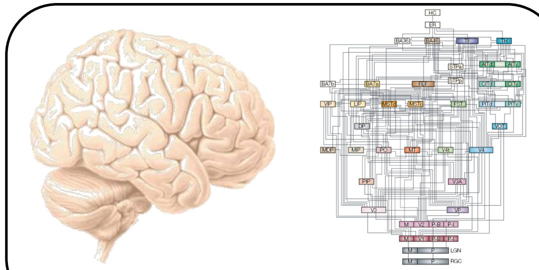
On the computational architecture of the neocortex
(D. Mumford 1992, *Biol Cybern*)

Predictive coding in the visual cortex
(Rao & Ballard 1999, *Nat Neurosci*)



Intellectual antecedents

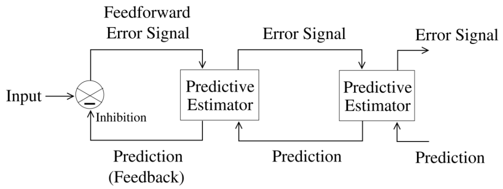
- Redundancy reduction



Felleman & Van Essen 1991, *Cereb Cortex*

Neuroanatomy

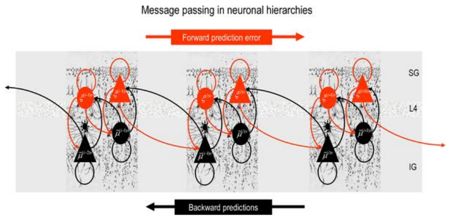
- Hierarchical organization of cortex



(Rao & Ballard 1999, *Nat Neurosci*)

Hierarchical PC model

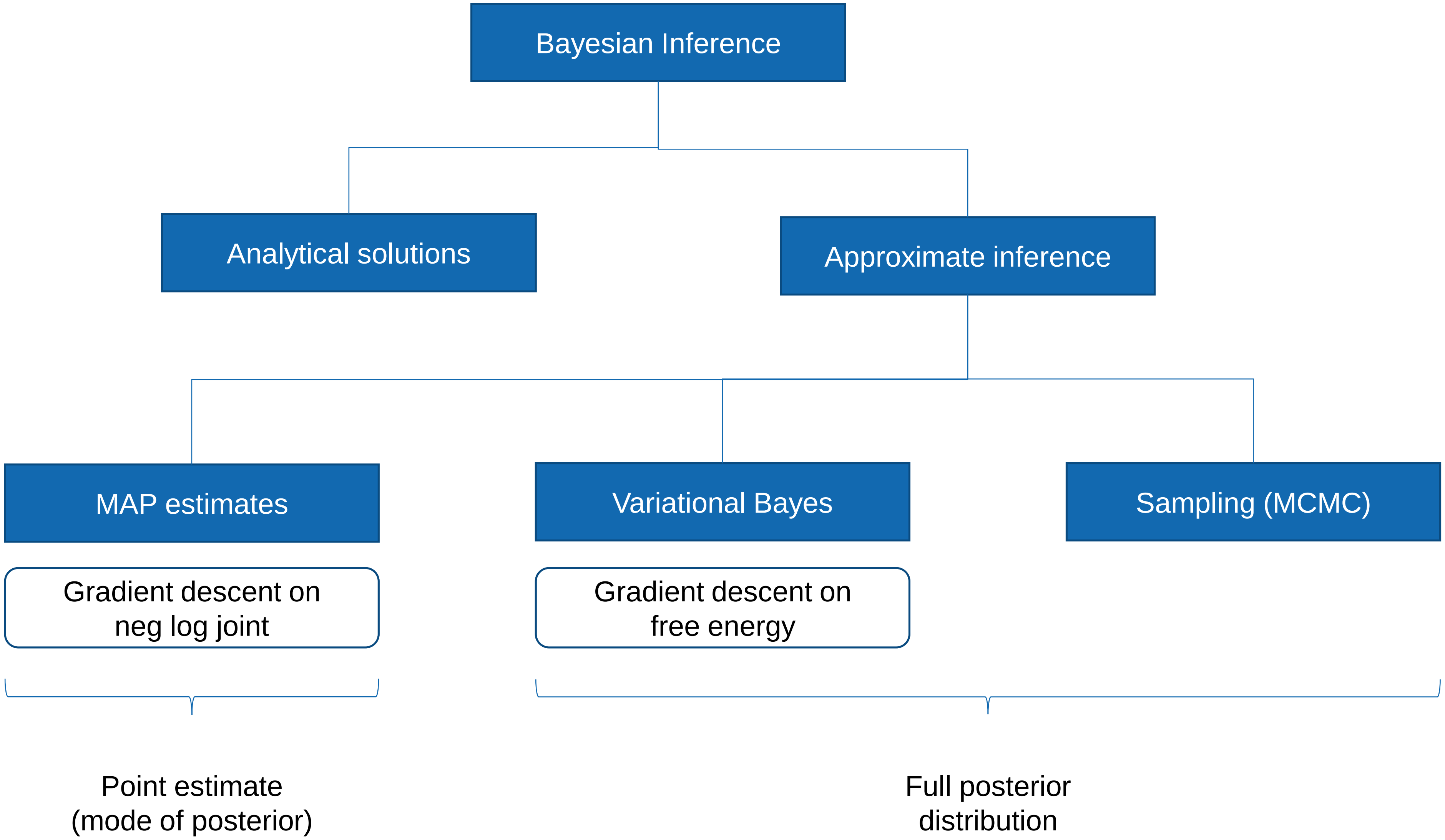
- Visual cortex
- Point estimate of posterior
- Static representations



Friston 2008, *PLoS Comput Biol*

PC as variational inference

Recap: Methods for Bayesian inference **Generative Models** lecture (Day 2, Klaas Enno Stephan)



Recap: Variational inference

VB & MCMC lecture (Day 2, Lionel Rigoux)

$$p(x|y,m) = \frac{p(y|x,m)p(x|m)}{p(y|m)} \qquad p(y|m) = \int p(y|x,m)p(x|m)dx$$

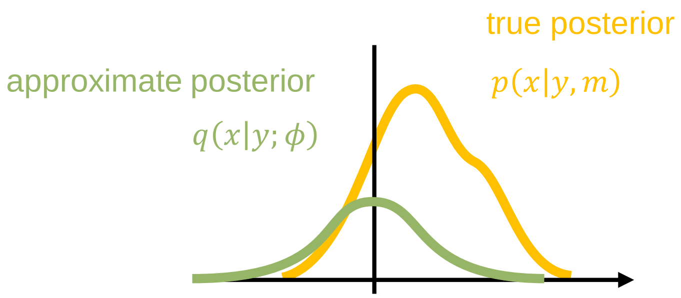


Figure adapted from slide by Yu Yao

Approximate posterior $q(x|y;\phi)$ e.g. for q Gaussian, $\phi = \{\mu, \Sigma\}$

Find best proxy $q^*(x|y;\phi) = argmin_\phi D_{KL}[q(x|y;\phi)||p(x|y,m)]$

Recap: Variational inference

VB & MCMC lecture (Day 2, Lionel Rigoux)

$$p(x|y,m) = \frac{p(y|x,m)p(x|m)}{p(y|m)} \qquad p(y|m) = \int p(y|x,m)p(x|m)dx$$

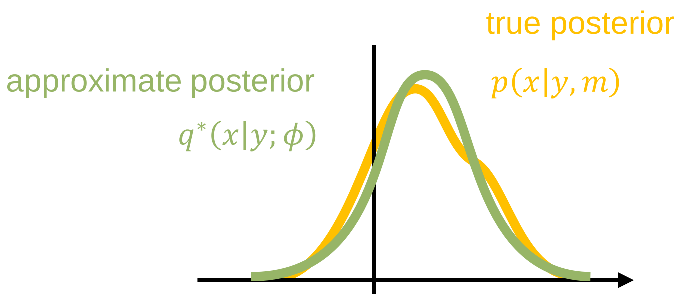


Figure adapted from slide by Yu Yao

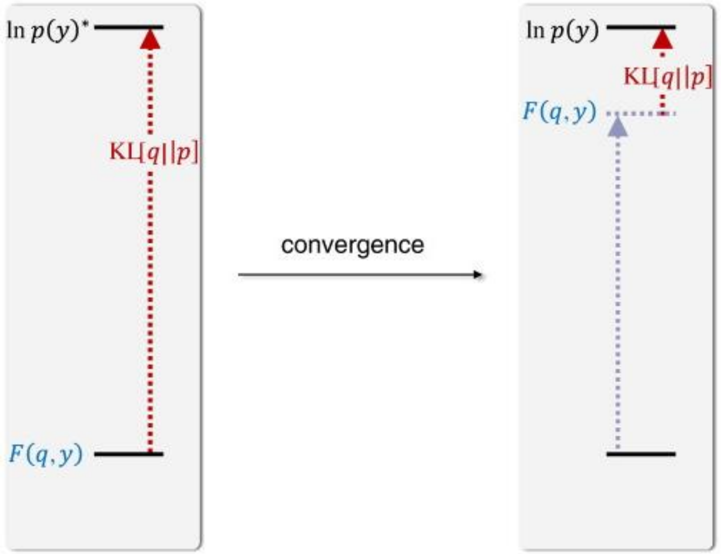
Approximate posterior $\quad q(x|y;\phi) \qquad$ e.g. for q Gaussian, $\phi = \{\mu, \Sigma\}$

Find best proxy $\quad q^*(x|y;\phi) = argmin_\phi D_{KL}[q(x|y;\phi)||p(x|y,m)]$

$$D_{KL}[q(x|y;\phi)||p(x|y,m)] = \ln p(y|m) - \int q(x|y;\phi)\,\frac{p(x,y|m)}{q(x|y;\phi)}dx$$

$$= \ln p(y|m) - F$$

$$\ln p(y|m) = D_{KL}[q(x|y;\phi)||p(x|y,m)] + F(q(x|y;\phi), p(x,y|m))$$



Stephan et al. 2017 *NeuroImage*

Predictive coding as variational inference

The free energy formulation of predictive coding

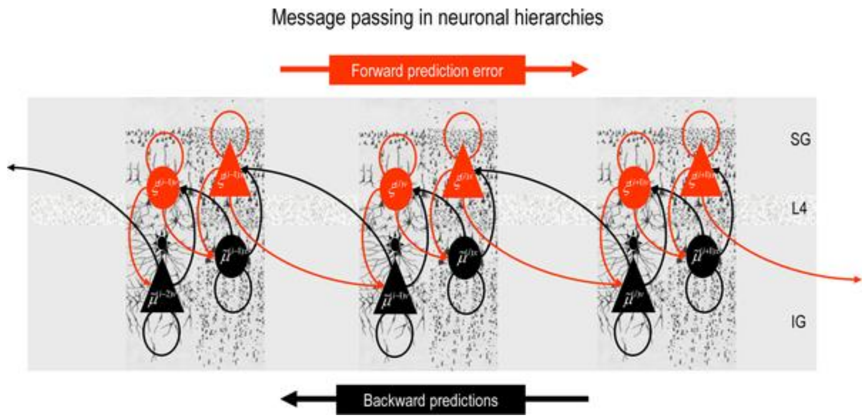
(Friston 2003, 2005, 2008)

- Minimal neuronal model
  - PE units (SG layers)
  - Prediction units (IG layers)

  at each level of the hierarchy

  ⇒ canonical microcircuit model (Bastos et al. 2012, *Neuron*)
- Model dynamics
  - Differential equations
  - Gradient descent on free energy F
- Importance of precision
- Extension to …
  - Temporal sequences (dynamic environment) ⇒ minimize free action $\bar{F}$
  - Action (active inference) (Friston et al. 2010, *Biol Cybern*; Adams et al. 2013, *Brain Struct Funct*)

Active Inference lecture
(Day 3, Thomas Parr)



Friston 2008, *PLoS Comput Biol*

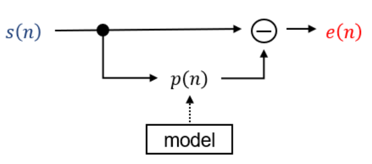
$$F = \int q(x|y;\phi)\,\frac{p(x,y|m)}{q(x|y;\phi)}\,dx$$

$$\bar{F} = \int F_t\,dt$$

Predictive Coding as neuroscientific theory

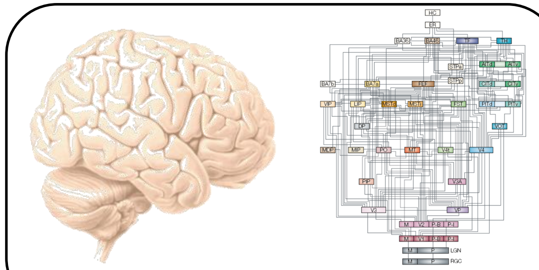
Non-Bayesian

On the computational architecture of the neocortex
(D. Mumford 1992, *Biol Cybern*)



Intellectual antecedents

- Redundancy reduction



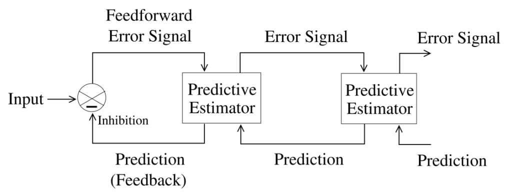
Felleman & Van Essen 1991, *Cereb Cortex*

Neuroanatomy

- Hierarchical organization of cortex

PC as approximate Bayesian inference

Predictive coding in the visual cortex
(Rao & Ballard 1999, *Nat Neurosci*)



(Rao & Ballard 1999, *Nat Neurosci*)

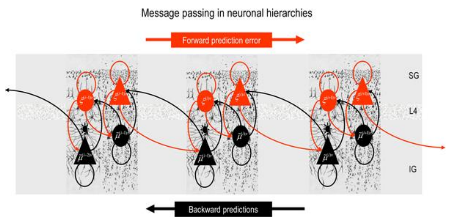
Hierarchical PC model

- Visual cortex
- Point estimate of posterior
- Static representations

Learning and Inference in the Brain
(Friston 2003, *Neural Netw*)

A theory of cortical responses
(Friston 2005, *Phil Trans Royal Soc B*)

Hierarchical Models in the Brain
(Friston 2008, *PLoS Comput Biol*)



Friston 2008, *PLoS Comput Biol*

PC as variational inference

- Cortical function
- Estimate full posterior
- Dynamic representations

Predictive coding in computational psychiatry

Predictive coding in computational psychiatry

The role of precision

- Finding the right balance
- Disorders of precision?

Schizophrenia lecture
(Day 1, Jakob Siemerkus)

Schizophrenia/Psychosis

(Stephan et al. 2006, *Biol Psychiatry*; Corlett et al. 2011, *NPP*; Adams et al. 2013, *Front Psychiatry*; Friston et al. 2016, *Schizophr Res*; Sterzer et al. 2018, *Biol Psychiatry*)

Autism lecture
(Day 1, Helene Haker Rössler)

Autism Spectrum Disorder

(Pellicano & Burr 2012, *TiCS*; Van de Cruys et al. 2014, *Psychol Rev*; Lawson et al. 2014, *Front Hum Neurosci*; Haker et al. 2016, *Front Psychiatry*; Lawson et al. 2017, *Nat Neurosci*)

From exteroception…

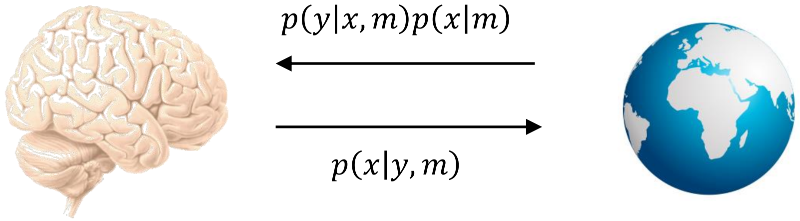


Figure adapted from a slide by Klaas Enno Stephan

Predictive coding in computational psychiatry

Schizophrenia lecture
(Day 1, Jakob Siemerkus)

Schizophrenia/Psychosis

(Stephan et al. 2006, *Biol Psychiatry*; Corlett et al. 2011, *NPP*; Adams et al. 2013, *Front Psychiatry*; Friston et al. 2016, *Schizophr Res*; Sterzer et al. 2018, *Biol Psychiatry*)

Autism lecture
(Day 1, Helene Haker Rössler)

Autism Spectrum Disorder

(Pellicano & Burr 2012, *TiCS*; Van de Cruys et al. 2014, *Psychol Rev*; Lawson et al. 2014, *Front Hum Neurosci*; Haker et al. 2016, *Front Psychiatry*; Lawson et al. 2017, *Nat Neurosci*)

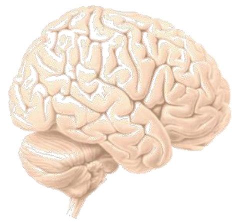
The role of precision

- Finding the right balance
- Disorders of precision?

From exteroception to interoception

- Interoceptive predictive coding
  (Seth et al. 2012, *Front Psychol*; Seth 2013, *TiCS;* Barrett & Simmons 2015, *Nature Rev Neurosci*;)
- Crucial role in mental health disorders
- Fatigue & depression (Stephan et al. 2016, *Front Hum Neurosci*)

Fatigue lecture
(Day 1, Inês Pereira)



$$\xleftarrow{p(y|x,m)p(x|m)}$$

$$\xrightarrow[p(x|y,m)]{}$$

Figure adapted from a slide by Klaas Enno Stephan

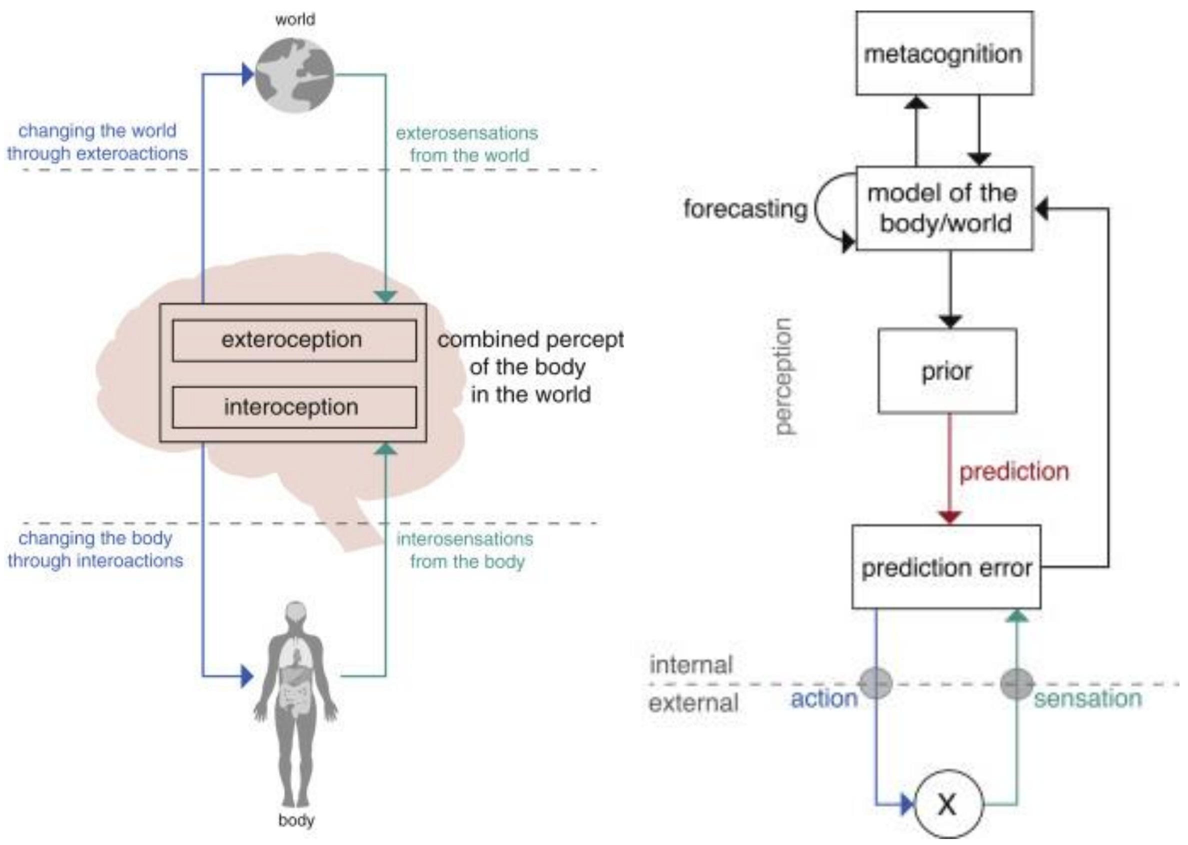
Hierarchical Bayesian Inference in Computational Psychiatry
(Petzschner et al. 2017, *Biol Psychiatry*)

- General framework to model adaptive behaviour

- Possible primary disruption at:
  - Sensory inputs (sensations)
  - Inference (perception)
  - Forecasting
  - Control (action)
  - Metacognition

- At any of the above, possible disturbance of:
  - predictions
  - prediction error computation
  - Estimation of precision

⇒ guide differential diagnosis



Petzschner et al. 2017, *Biol Psychiatry*

Hierarchical Bayesian Inference in Computational Psychiatry
(Petzschner et al. 2017, *Biol Psychiatry*)

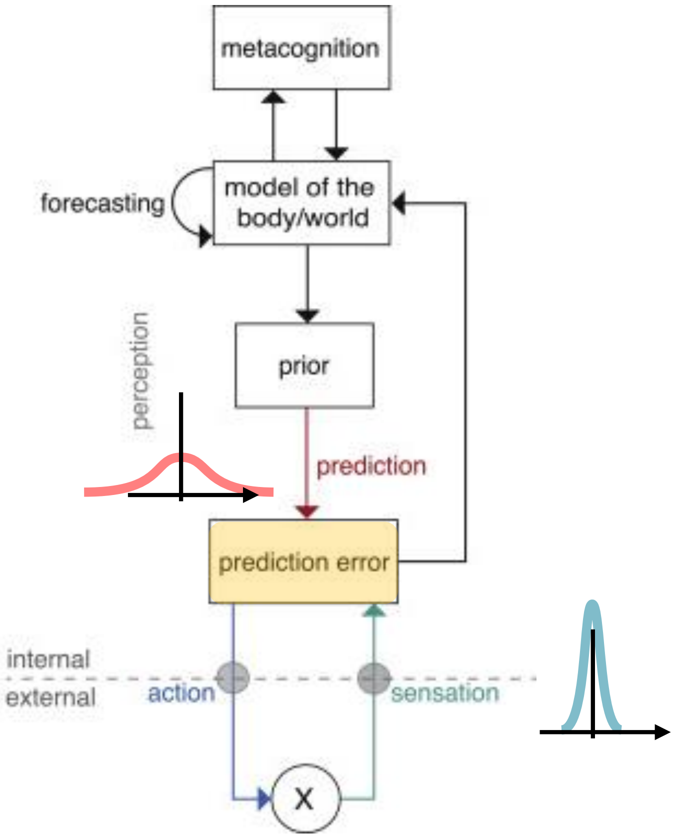
Example: Autism Spectrum Disorder

- Patients: excessive processing of irrelevant details

- 2 competing explanations
  - Sensory inputs of overwhelming precision
  - Too imprecise higher-order beliefs

  ⇒ large PEs during perception

- Disambiguate 2 hypotheses:
  - Assess individual sensory processing (experiment + model)
  - Detect (sub)groups

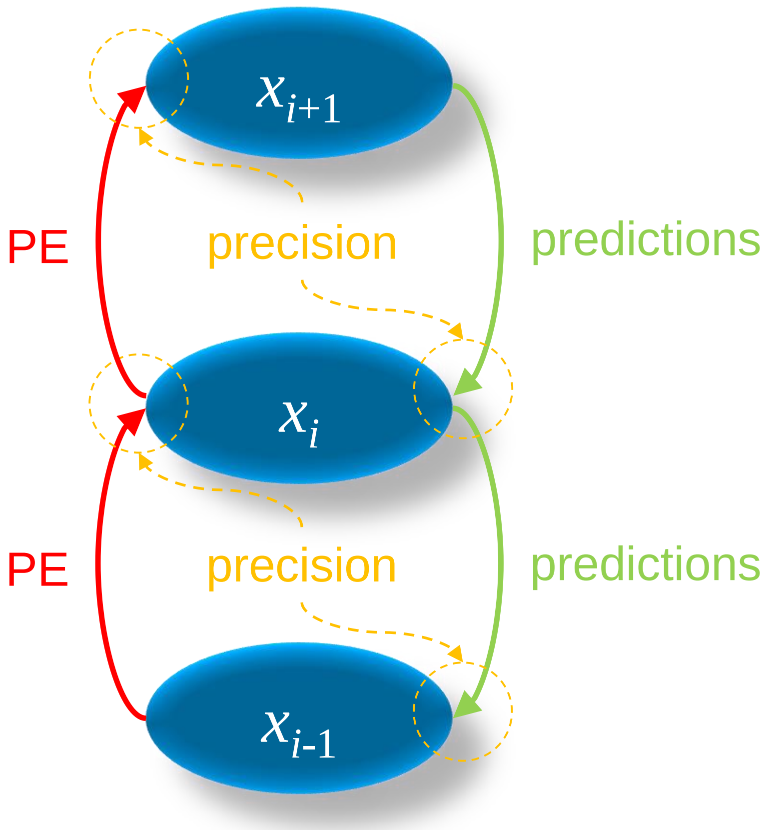


Petzschner et al. 2017, *Biol Psychiatry*

Predictive coding in a nutshell

$$\Delta\text{belief} \sim \text{precision} \times \text{PE}$$

- Possible way of implementing Hierarchical Bayesian inference in the brain
- Based on
  - Redundancy reduction
  - Hierarchical organization of cortex
- Computational quantities:
  - Each layer makes **predictions** about activity in layer immediately below
  - Predictions are compared with inputs of each layer
  - **Prediction errors (PE)** signalled upwards
  - Relative influence of PEs and predictions is determined by their relative **precision** (certainty)
- Goal of the brain:
  - minimize PE at each level of the hierarchy
- Utility of this framework for Computational Psychiatry & Computational Psychosomatics



Adapted from Stephan et al. 2016, *Brain*

Further reading

REVIEWS

Theoretical & experimental review Millidge et al. 2021, *arXiv:2107.12979*

Experimental evidence for PC in the brain Walsh et al. 2020, *Ann N Y Acad Sci*

PC algorithms Spratling et al. 2017, *Brain Cogn*

TUTORIALS

PC as variational inference Bogacz 2017, *J Math Psychol*; Buckley 2017, *J Math Psychol*

OTHER

PC & laminar fMRI Stephan et al. 2019, *NeuroImage*

PC networks and backpropagation of error algorithm Whittington & Bogacz 2017, *Neural Comput*; Song et al. 2020, *Adv Neural Inf Process Syst*

PC, variational autoencoders & normalizing flows Marino 2020, *arXiv:2011.07464*

Universität Zürich[UZH] ETH zürich

Translational Neuromodeling Unit

Thank you!

Lilian Weber
Matthias Müller-Schrader
Sandra Iglesias
Stefan Frässle
Klaas Enno Stephan

Alex Hess

E-Mail: hess@biomed.ee.ethz.ch

@alex_j_hess