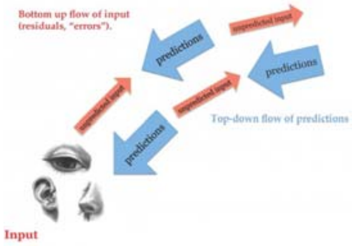
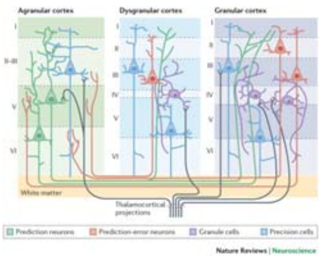


Introduction & computational model

Predictive Processing Research Seminar

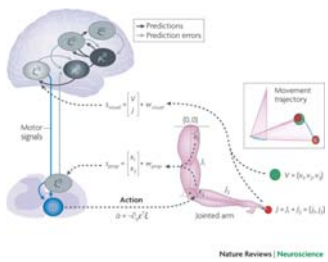
Introduction & computational model

Predictive Processing Research Seminar



Introduction & computational model

Predictive Processing Research Seminar

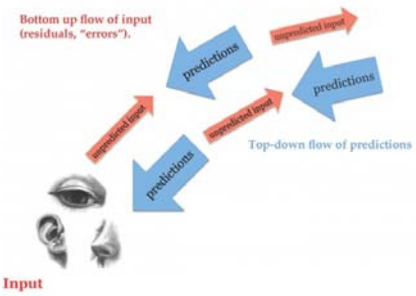


Introduction & computational model

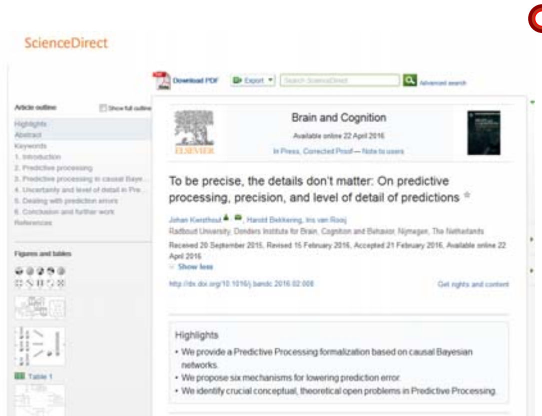
Predictive Processing Research Seminar

A novel view of the brain

"…essentially a sophisticated hypothesis-testing mechanism…"



Clark, A. (2015). Surfing Uncertainty.



Predictive Processing vs. Predictive Coding



- Not **Friston**-ese, but **Clark**-onian / **Hohwy**-esque
- Focus on conceptual principle rather than Friston's "free-energy driven predictive coding" which is closely tied to the cortical hierarchy – we wish to abstract away from that and stay at Marr's computational level
- Computational translation of conceptual principles

Predictive Processing & Marr's levels

Computational level Conceptual description	**Predictive Processing** [Clark, Hohwy] *Keywords:* predictions at various levels of detail, precision-weighted prediction errors, hypothesis updating, model revision
Algorithmic level Process-level description	**Predictive Coding** [Friston, Kilner, Spratling, Mathys] *Keywords:* low-level cognition, continuous Gaussian models, variational Bayes approximations, free-energy principle
Implementational level Neuronal level description	**Cortical microcircuits** [Bastos] *Keywords:* pyramid cells, feedforward-feedback connections, canonical microcircuits

Predictive Processing & Marr's levels

Computational level Conceptual description	**Predictive Processing** [Clark, Hohwy] *Keywords:* predictions at various levels of detail, precision-weighted prediction errors, hypothesis updating, model revision
Algorithmic level Process-level description	**Belief Propagation** [Sanborn, Griffiths, Tenenbaum] *Keywords:* high-level cognition, structured discrete graphical models, sampling approximations, particle filtering
Implementational level Neuronal level description	**Networks of Spiking Neurons** [Maass, Lengyel, Hennequin] *Keywords:* Boltzmann machines, switching rate, noisy spikes

Predictive Processing

Brain as **prediction machine**
- *The brain continuously makes predictions about future sensory evidence based on its current best model of the causes of such evidence*

Bayesian Brain
- *The brain combines prior knowledge with sensory evidence (from various sources) in a Bayesian way*

Hierarchical Brain
- *The brain is organized in a hierarchical way, where "high level" information influences "low level" information and vice versa*

Key sub-processes

- Making **predictions** of expected input based on the generative models
- Comparing predicted inputs with actual inputs and **computing prediction error**
- **Explaining away** prediction errors (minimizing prediction error)
- **Learning** and adapting generative models

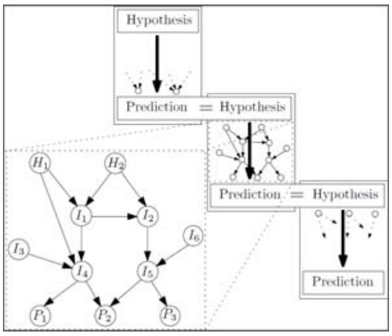
From conceptual idea to formal model

- Predictive processing is assumed to **explain and unify all of cognition**, including higher cognition
- To model, e.g., complex social interactions, Theory of Mind, intention recognition, and problem solving, we need rich enough knowledge structures to model dependences

→ We argue (Otworowska et al., 2014) that simple Gaussian models are **not sufficiently rich** models for higher cognition

→ We propose to use **causal Bayesian networks** as knowledge structures instead to describe predictive processing

Computational model – prediction generation



Hypothesis variables *Hyp* = {H_1, H_2}

Prediction variables *Pred* = {P_1, P_2}

Intermediate variables *Int* = {I_1, … I_6}

Computational model – error estimation

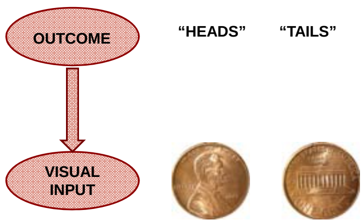
- Prediction and Observation are **probability distributions** over the prediction variables *Pred*
- Prediction is defined as computing the **posterior distribution** $Pr_{(Pred)}$ given the parameters in the network
- Prediction error is set difference $Pr_{(obs)}$ - $Pr_{(Pred)}$
- Estimating the size of this error is defined as computing a KL- divergence or **relative entropy** between predicted distribution and observed distribution

$$D_{KL}(Pr_{(Obs)} \| Pr_{(Pred)}) = \sum_{p \in \Omega(Pred)} Pr_{Obs}(\mathbf{p}) \log \left(\frac{Pr_{Obs}(\mathbf{p})}{Pr_{Pred}(\mathbf{p})} \right)$$

Precision of prediction & prediction error

- **Precision** is a property of both a **prediction** and of a **prediction error**
- **Precision of a prediction** describes how much uncertainty there is in a prediction (and consequently, how informative the actual observation of what was predicted will be)
- **Precision of a prediction error** describes what proportion of this uncertainty can be attributed to inherent stochastic nature of the process that caused the outcome of the prediction → **precision-weighted prediction errors**

Example: tossing coins



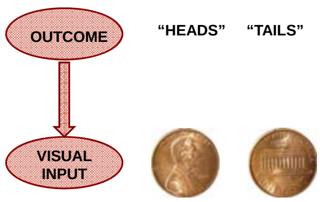
On the whiteboard!

Computational model – error minimization

- Prediction error minimization: “doing something” such that D_{KL}(*Obs* || *Pred*) is minimized
- Four possible ways of “doing something” (Kwisthout, van Rooij, & Bekkering, 2017):
  - Belief revision (revise hyp probability distribution)
  - Model revision (revise parameters in the CPTs)
  - Passive intervention (evidence gathering)
  - Active intervention (acting, i.e., setting variables)
- Each of them with the goal of lowering relative entropy
- We’ll go through them

Intuitive examples of lowering prediction error

- **Belief revision** – in cases with ‘expected uncertainty’ where the world model is stable but there is information carried by the prediction error
- On the whiteboard!



Intuitive examples of lowering prediction error

- **Model revision** – in cases with ‘unexpected uncertainty’ where the world model suddenly turns out to be misinformative
- Prediction errors can be dealt with by changing some of the **model parameters** (tuning) such that the model can better predict the observations
- E.g., your model about what constitutes a friendly greeting may need updating (for the 30+-people amongst us)

Intuitive examples of lowering prediction error

- **Passive intervention** – reduce prediction error by reducing uncertainty in the world: **add** additional observations
- This is what we intuitively do when confronted with the “train effect”: when you’re sitting in a train that is standing still at the station and you are looking at an opposite train – who is moving?
- You’d probably look at a stationary point to reduce uncertainty (e.g., the railway station buildings)

Intuitive examples of lowering prediction error

- **Active intervention** – reduce prediction error by actively intervening in the world (active inference): bring prediction and observation closer together by **changing** the observation
- This has been proposed as a means of coupling action and perception in a single framework, where motor acts are the result of a mismatch between a “predicted” (expected) state and the actual, perceived state of the world

Next time: how to develop a generative model?