Leaving Andy Clark's 'safe shores': Scaling predictive processing to higher cognition

Johan Kwisthout, Maria Otworowska, Harold Bekkering & Iris van Rooij

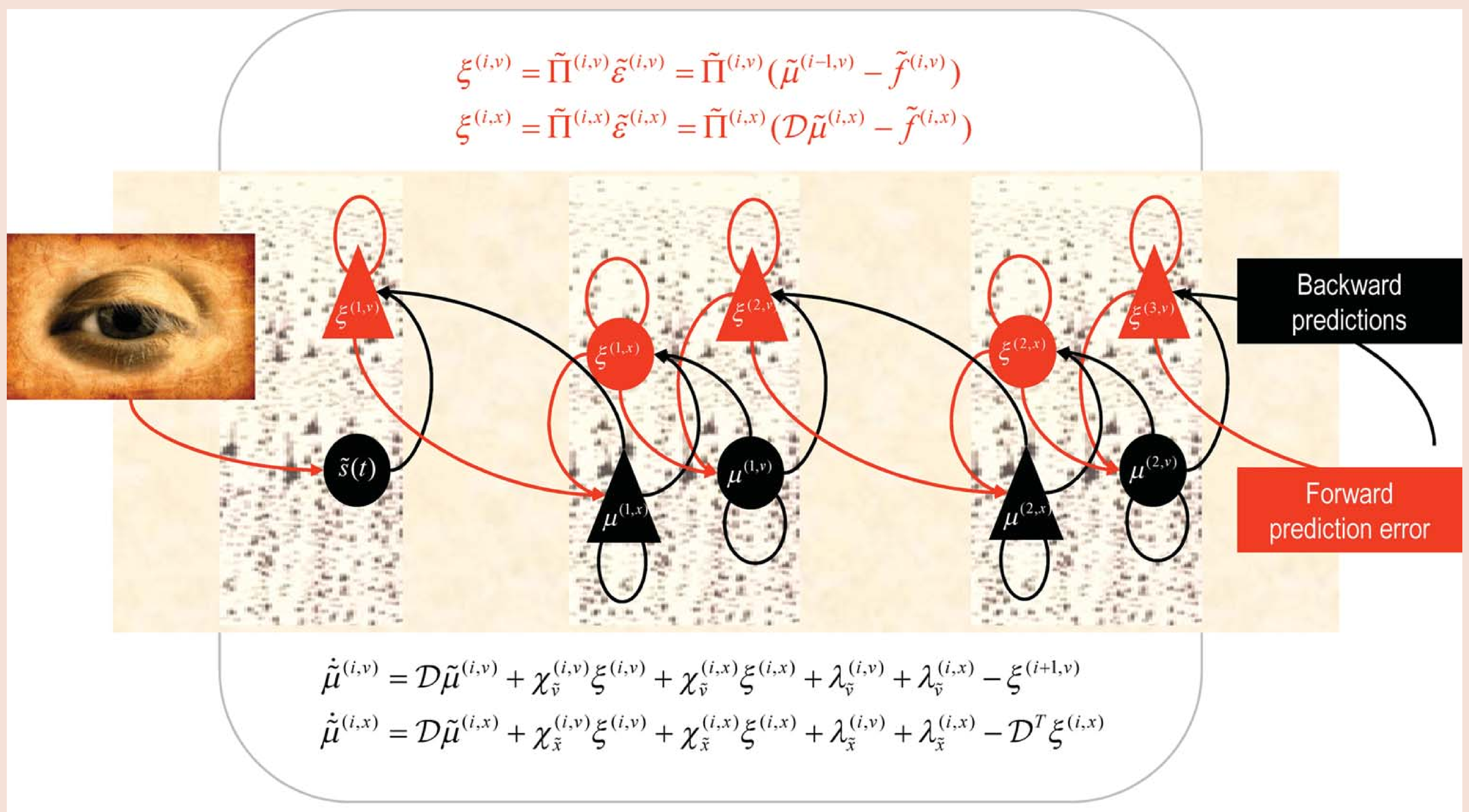
Donders Institute for brain, cognition and behaviour

Radboud University Nijmegen

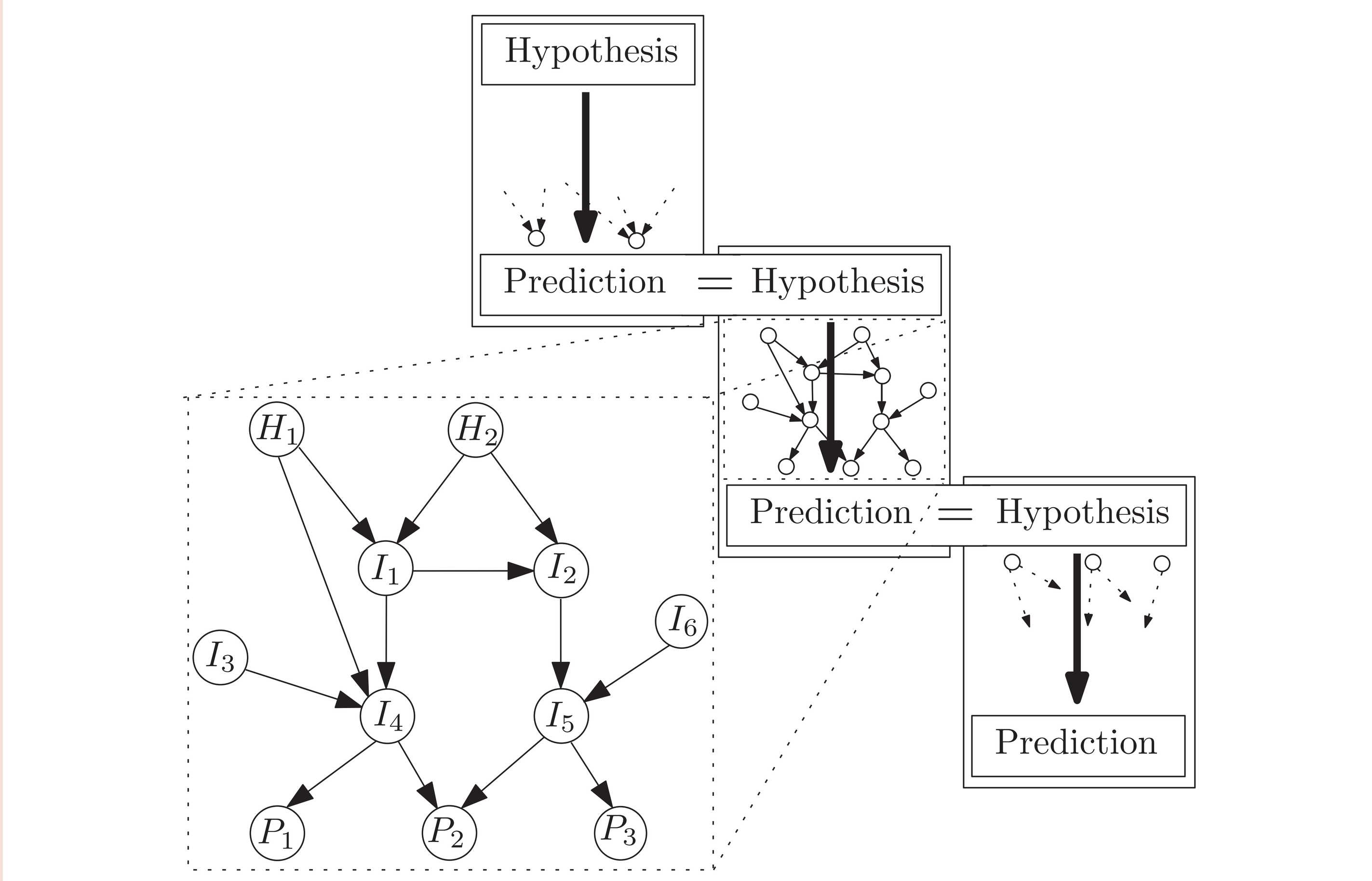
Introduction

The predictive processing principle has gained considerable interest as a strong candidate for explaining the unifying principles that underlie the brain's activity, spanning the entire range of cortical activity from perception and motor control to social cognition and theory of mind. Yet, the current mathematical framework that supports predictive processing heavily relies on the Laplace assumption, i.e., stochastic dependencies are assumed to be (multivariate) Gaussian densities. This assumption becomes problematic when scaling the predictive processing principle to higher cognition, that is, when we "depart further and further from the safe shores of basic perception and motor control?" (Clark, 2013). In order to be scalable to higher cognition, the mathematical framework underlying predictive processing should be able to represent structured, unordered, discontinuous, and non-monotone information and relations. We introduce a formulation of predictive processing in terms of structured Bayesian networks that overcomes the limitations of Gaussian densities.

Key Question: does it scale?



Typical example of predictive coding principle in the brain (Feldman & Friston, 2010)



Predictive processing cast in Bayesian network models (Kwisthout, 2014)

Defining predictive processing in terms of structured Bayesian network models (rather than simple Gaussian relations between hypothesis and prediction) allows us to extend the scope of the principle to social cognition with its complex, context-specific, non-linear and un-ordered dependencies. However, in order to provide a full mathematical model one needs to deal with the following issues: How to define precision, how to minimize prediction errors, and at what level of detail does knowledge need to be represented.

Precision

Problem: Precision is defined in predictive coding as the inverse of variance of a Gaussian distribution and as such tightly coupled to Gaussian uncertainty. There is no such corresponding notion in structured probability distributions.

Solution: Variance is a derived measure of the more general measure "entropy" of a distribution. The entropy of both Gaussian and structured distributions represents the uncertainty associated with this distribution.

Error minimization

Problem: How can we define the mechanisms for lowering prediction error by active inference, hypothesis updating etc. in structured probability distributions?

Solution: Bayesian networks allow for elegant formalizations of model revision, hypothesis revision, intervention (active inference) and sampling the environment (obtaining additional evidence). These error minimization strategies can be studied both independently and in interaction with each other.

Level of detail

Problem: When representing knowledge using structured models, we need to decide on the level of detail of the models.

Solution: Such representations actually allow us to disentangle "uncertainty" and "level of detail", both of which are confounded in the Gaussian notion of precision. The active modulating of the level of detail of predictions and observations may help to explain phenomena such as attention, alertness, and learning of generative models.

References

Blokpoel, M., Kwisthout, J., and van Rooij, I. (2012). When can predictive brains be truly Bayesian? Commentary to: Clark (2013). Whatever next? Neural prediction, situated agents, and the future of cognitive science. *Frontiers in Theoretical and Philosophical Psychology*, 3 (406

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181-204.

Feldman, H. and Friston, K.J. (2010). Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience*, 4 (215).

Friston, K.J., Adams, R.A., Perrinet, L., and Breakspear, M. (2012). Perceptions as hypotheses: saccades as experiments. *Frontiers in Psychology*, 3(151).

Heil, L., van Pelt, S., Kwisthout, J., van Rooij, I., and Bekkering, H. (2013). Higher-level processes in the formation and application of associations during action understanding. Commentary to: Cook, Bird, Catmur, Press, and Heyes (2014). Mirror neurons: From origin to function. *Behavioral and Brain Sciences*, 37(2), 202 - 203.

Kwisthout, J. (2014). Minimizing relative entropy in Hierarchical Predictive Coding. Accepted to The Seventh European Workshop on Probabilistic Graphical Models.

Otworowska, M.E., Kwisthout, J., and van Rooij, I. (2014). Counter-factual mathematics of counterfactual predictive models. *Frontiers in Consciencess Research* 5 (801).

www.dcc.nl/~johank

j.kwisthout@donders.ru.nl