

Chapter 1

Introduction

1.1 Sensory predictions across space and time

The brain is often framed as a general purpose “prediction machine” (Hawkins & Blakeslee, 2004; Clark, 2013). In this framework, the sole evolved function of the neocortex is to minimize error in its representation of predictions about the physical world. This distillation of function is central to a number of models of neocortical function (e.g., Dayan, Hinton, Neal, & Zemel, 1995; Rao & Ballard, 1999; Lee & Mumford, 2003; Friston, 2005; George & Hawkins, 2009), but is surprisingly often overlooked in psychology and neuroscience investigations of sensory processing. For example, most experiments are designed to measure evoked responses to a randomly chosen, isolated stimulus under the tacit assumption that response variability is irrelevant noise that should be averaged out across many presentations. Computational models of perceptual processing often operate under similar assumptions in which stimuli are presented as random “snapshots” from which some common set of features should be learned to minimize representational variability across presentations (e.g., Riesenhuber & Poggio, 1999; Serre, Oliva, & Poggio, 2007; O’Reilly, Wyatte, Herd, Mingus, & Jilk, 2013; although see Foldiak, 1991, for a notable exception). These

experimental and modeling assumptions stand in contrast to the event structure of the physical world, which is highly structured from one moment to the next. It could be the case that response variability is not simply due to noise, but is related to meaningful predictive processing that captures this spatiotemporal structure (Arieli, Sterkin, Grinvald, & Aertsen, 1996; Wilder, Jones, Ahmed, Curran, & Mozer, 2013; Fischer & Whitney, 2014).

There are a number of important questions that need answered to fully characterize prediction and its role in sensory processing. What are the mechanisms responsible for making predictions? Computationally, there is a fundamental tradeoff in making decisions about and generating actions from the constant stream sensory information to versus actively generating predictions about what will happen next. Do standard mechanisms balance these tradeoffs or is there special purpose, dissociable machinery specifically for predictive processing. Another line of questioning is concerned with how the brain knows *when* to make predictions. Prediction requires integrating information over some time frame and using the result to drive the actual prediction, but when should integration start? And for how long?

The goal of this thesis is to develop a line of research designed to provide answers to some of these questions and raise others. The work is largely predicated on a modeling framework referred to as *LeabraTI* (Temporal Integration), an extension of standard Leabra cortical learning algorithms (O'Reilly & Munakata, 2000; O'Reilly, Munakata, Frank, Hazy, & Contributors, 2012) that describes how prediction is accomplished in biological neural circuits. The framework brings together a large number of independent findings from the systems neuroscience literature to describe exactly how multiple interacting mechanisms trade off prediction with sensory processing and learn associations across temporally extended sequences of input.

The fundamental proposal of LeabraTI is that time in the brain is discretized into roughly 100 ms intervals that allow interleaving predictions and sensory processing over time. These intervals correspond to individual cycles of the widely observed alpha rhythm over posterior cortical areas (Klimesch, Sauseng, & Hanslmayr, 2007; Mathewson et al., 2011). The temporal interleaving of prediction and sensory processing allows standard error-driven learning to operate

TODO

1.2 Brief review of spatial- and temporal-related processing

1.3 Thesis organization

References

- Arieli, A., Sterkin, A., Grinvald, A., & Aertsen, A. (1996). Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. *Science*, *273*(5283), 1868–1871.
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204.
- Dayan, P., Hinton, G. E., Neal, R. N., & Zemel, R. S. (1995). The Helmholtz machine. *Neural Computation*, *7*(5), 889–904.
- Fischer, J., & Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience*, *17*(5), 738–743.
- Foldiak, P. (1991). Learning invariance from transformation sequences. *Neural Computation*, *3*(2), 194–200.
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London*, *360*(1456), 815–836.
- George, D., & Hawkins, J. (2009). Towards a mathematical theory of cortical micro-circuits. *PLoS Computational Biology*, *5*(10).
- Hawkins, J., & Blakeslee, S. (2004). *On Intelligence*. New York, NY: Times Books.
- Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition-timing hypothesis. *Brain Research Reviews*, *53*(1), 63–88.
- Lee, T. S., & Mumford, D. (2003). Hierarchical bayesian inference in the visual cortex. *Journal of the Optical Society of America*, *20*(7), 1434–1448.
- Mathewson, K. E., Lleras, A., Beck, D. M., Fabiani, M., Ro, T., & Gratton, G. (2011). Pulsed out of awareness: EEG alpha oscillations represent a pulsed-inhibition of ongoing cortical processing. *Frontiers in Psychology*, *2*.
- O'Reilly, R. C., & Munakata, Y. (2000). *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain*. Cambridge, MA: The MIT Press.
- O'Reilly, R. C., Munakata, Y., Frank, M. J., Hazy, T. E., & Contributors (2012). *Computational Cognitive Neuroscience*. Wiki Book, 1st Edition, URL: <http://ccnbook.colorado.edu>.
- O'Reilly, R. C., Wyatte, D., Herd, S., Mingus, B., & Jilk, D. J. (2013). Recurrent processing during object recognition. *Frontiers in Psychology*, *4*, 124.

- Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. Nature Neuroscience, 2(1), 79–87.
- Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. Nature Neuroscience, 2(11), 1019–1025.
- Serre, T., Oliva, A., & Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. Proceedings of the National Academy of Sciences of the United States of America, 104(15), 6424–6429.
- Wilder, M. H., Jones, M., Ahmed, A. A., Curran, T., & Mozer, M. C. (2013). The persistent impact of incidental experience. Psychonomic Bulletin & Review, 20(6), 1221–1231.