Chapter 1

Introduction

1.1 Sensory predictions and temporal integration

The brain is often framed as a general purpose "prediction machine" (HawkinsBlakeslee04Clark13). The fundamental assertion of this framework is that 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., DayanHintonNealEtAl95RaoBallard99LeeMumford03Friston05GeorgeHawkins09), 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 averages 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., Fukushima80RiesenhuberPoggio99MasquelierThorpe07OReillyWyatteHerdEtAl13; although see Foldiak91, 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 does not simply reflect noise, but is actually related to meaningful predictive processing that captures this temporal structure (ArieliSterkinGrinvaldEtAl96WilderJonesAhmedEtAl13FischerWhitney14).

There are a number of important questions that need answered to fully characterize prediction and its role in sensory processing. What are the neural mechanisms responsible for making predictions? Computationally, there is a fundamental tradeoff in making decisions about and generating actions from the constant stream sensory information 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 how long should it last?

The goal of this thesis is to develop a line of research designed to provide answers to some of these questions and of course, to raise others. The work is largely predicated on a modeling framework referred to as LeabraTI (TI: Temporal Integration), an extension of the standard Leabra cortical learning algorithm (OReillyMunakata00OReillyMunakataFrankEtAl12) 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 biological details of LeabraTI give rise to a number of testable predictions that can be used to determine the validity of the overall framework. Central to the these testable predictions is the idea that internally generated predictions and sensory events are interleaved through the same neural tissue over intervals of 100 ms. These intervals correspond to individual cycles of the widely observed $\sim$10 Hz alpha rhythm over posterior cortical areas (PalvaPalva07HanslmayrGrossKlimeschEtAl11VanRullenBuschDrewesEtAl11). The temporal interleaving of prediction and sensory processing allows powerful error-driven learning mechanisms to minimize prediction error over multiple episodes, but with the side-effect of discretization artifacts and other temporal oddities due to suppressing sensory processing in favor of prediction for a portion of each 100 ms period.

The empirical work described in this thesis takes advantage of the brain's putative 10 Hz prediction-sensation rate by presenting exogenous stimulation either in phase or out of phase with this endogenous processing. This allows testing of how the spatiotemporal predictability of stimuli influence their encoding for perceptual judgements or prolonged learning. The thesis also describes a neural network model that implements of the broader LeabraTI framework with the goal of accounting for the results of the experimental work.

1.2 Organization of the thesis

The organization of this thesis is as follows. Chapter **??** contains a comprehensive description of the LeabraTI framework in terms of the low-level biological details of the cortical microcircuitry and response properties required for the temporally interleaved prediction-sensation computation. The chapter also compares LeabraTI with other modeling frameworks and describes a number of testable predictions that differentiate it and can more generally be used to determine its overall validity.

Chapters **??** and **??** describe two experiments designed to test the fundamental predictions of the LeabraTI. The Chapter **??** experiment involves manipulating the predictability of time- and space- varying stimulus sequences so that they are either presented in phase or out of phase with the endogenous $\sim$10 Hz alpha rhythm. The stimulus sequences are followed by a perceptual judgement to determine the effect of multiple successful (or unsuccessful) sensory predictions on the encoded representation. EEG is also recorded during the experiment to investigate the effects of these manipulations on endogenous alpha oscillations. The Chapter **??** experiment expands on the basic Chapter **??** experimental paradigm to investigate the behavioral effects of predictability on prolonged learning of a subset of the stimuli.

Chapter **??** describes the neural network model that implements the broader LeabraTI framework which is capable of reproducing the results of the Chapter **??** and **??** experiments. The model also provides insight into how predictive learning can alter the representation of the stimuli over prolonged periods of exposure. Finally, Chapter **??** discusses the results of the cumulative work accompanied by several open questions and directions for future work.