Neural Circuit Discovery

Eric Jonas
Konrad Kording

November 4, 2013

Abstract

The brain is made out of multiple types of neurons, and a neurons type affects both its behavior and which other neurons it connects to. Emerging techniques (connectomics) allow measuring the connectivity matrix between many neurons. Estimating types from connectomics is difficult as distance is usually more important than type. Here we describe a non-parametric Bayesian technique that overcomes this problem and discovers neuron types based on connections only. We show that the approach recovers known neuron types in the retina, reveals interesting structure in the nervous system of c. elegans, and automatically discovers the structure of microprocessors. Extracting meaningful structure from connectivity data promises to enable new experiments and to constrain theories of brain function.

1 Introduction

Computing systems, biological or human-engineered, contain computing elements that can be classified into "type", which give insight into function.

The advent of connectomics data allows discovery of cell types based on connectivity information.

Here we describe a Bayesian non-parametric model that can discover both the cell types and their patterns of interconnection automatically from connectomics data.

We apply it to three very-different computing systems: recently-released mouse retina connectome, the c. elegans connectome, and a "connectome" obtained by reverse-engineering a portion of a classical microprocessor.

In all cases, we recover known types and suggest possible new types.

2 Methods

Stochastic block models assume a hidden or "latent" type is associated wtih each node that ultimately influences its connectivity.

We extend this class of model to incorporate notions of spatial locality, as the connectivity of many neural systems is substantially constrained.

We perform posterior inference in this model using a series of Markov-chain Monte Carlo techniques.

3 Results

3.1 Synthetic Data

We show our model works by generating synthetic data with known spatial/connectivity patterns and recovering ground truth, even when the generating process makes assumptions very different from our model.

3.1.1 Mouse Retina

When applied to the mouse retina dataset, we recover spatially-homogeneous patterns of activity that reflect a coarsening of ground truth.

3.2 C. elegans

When applied to c. elegans, we segregate "head and sensor" interneurons from the spatially-distributed neurons along the body axis.

3.3 Microprocessor

We recover the logical structure of the three primary registers of the MOS6502 integrated circuit.

4 Discussion

We present a machine learning technique that allows cell types to be disocvered using only connectivity data.

The rise of clustering had a substantial impoact on molecular biology, and is now the only way that molecular biologists are able to deal with high-throughput sequencing technologies.

In the future, we plan to extend our model with the ability to handle other classes of information, including genetic and histochemical metadata.