Machine Intelligence from Cortical Networks (MICrONS) workshop

On Statistical Inference for the Cortical Column Conjecture

Carey E. Priebe

Department of Applied Mathematics & Statistics Johns Hopkins University, Baltimore, MD, USA

February 11-12, 2014

Outline

Introduction

Cortical Column Conjecture Random Graph Model Hypotheses Biologically Motivated Example

Cortical Column Conjecture Test Known Graph Alignment Effect of Graph Matching

Identifying Large Scale Structures

Effect of Misspecified Subgraphs Effect of Misspecified Subgraphs in Practice

Discussion

Cortical Column Conjecture Test

Identifying Large Scale Structures

Discussion

Motivation

Many contemporary theories of neural information processing suggest that the neocortex employs algorithms composed of repeated instances of a limited set of computing primitives.

There is a recognized need for tools for interrogating the structure of the cortical microcircuits believed to embody these primitives.

Cortical Column Conjecture

Neurons are connected in a graph that exhibits motifs representing <u>repeated</u> processing modules (Horton & Adams, 2005; Mountcastle, 1997).

Question

• Can we test whether an observed cortical graph satisfies the conjecture?

• Can we estimate relevant cortical computing parameters from an observed cortical graph?

Goal

To present a notional demonstration of how statistical inference on graphs can inform our understanding of cortical computing.



J. C. Horton and D. L. Adams. "The cortical column: a structrue without a function," *Philosophical transactions of the royal society B*, 360(1456):837-862, Apr. 2005.



V. Mountcastle. "The columnar organization of the neocortex," Brain, 120(4):701-722, Apr. 1997.





Pipeline

Cortical Column Conjecture Test

Identifying Large Scale Structures 00000 Discussion

Cortical Column Conjecture



 Cortical Column Conjecture Test

Identifying Large Scale Structures

Discussion

Cortical Column Graph



- cortical graph G: hierarchical block model on n vertices (neurons).
- induced subgraph H_r : stochastic block models on $n_r, r = 1, ..., R$.

- 1. identify large-scale structures in $G \rightarrow H_r$,
- 2. test if \hat{H}_r 's correspond to repeated motifs.

Cortical Column Conjecture Test

Identifying Large Scale Structures

Discussion

Stochastic Block Model

Let \mathcal{B}_K be the collection of symmetric $K \times K$ block-probability matrices;

$$\mathcal{B}_K := \{B \in [0, 1]^{K \times K} | B^T = B\}.$$

Let $\mathcal{M}_{n,K}$ be the collection of length K non-negative integer-valued vectors \vec{n} with entries summing to n;

$$\mathcal{M}_{n,K} := \{ \vec{n} \in \mathbb{N}^K | \sum_{k=1}^K \vec{n}_k = n \}.$$

Given positive integers n and K, consider $\vec{n} \in \mathcal{M}_{n,K}$ and $B \in \mathcal{B}_K$. We say a random graph G on vertex set V is a stochastic block model graph $G \sim SBM(V, \vec{n}, B)$ if the vertices V are partitioned into subsets V_1, \cdots, V_K with sizes given by $\vec{n} = (\vec{n}_1, \cdots, \vec{n}_K)$ and edges $\mathbb{1}\{u \sim v\}$ are independent Bernoulli random variables with $P[u \sim v] = B_{ij}$ for $u \in V_i$ and $v \in V_j$.

.

Hierarchical Stochastic Block Model

Given G = (V, E), where V = [n],

- $\{V_r\}_{r=1}^R$: a disjoint partition of the vertices V, and for each r = 1, 2, ..., R, write $n_r := |V_r|$.
- $H_r = \Omega(V_r) = (V_r, E_r).$
- $H_r \sim SBM(V_r, \vec{m}_r, B_r)$, where $\vec{m}_r \in \mathcal{M}_{n_r, K}$ and $B_r \in \mathcal{B}_K$.
- for $u \in V_r$ and $v \in V_{r'}$ with $r \neq r'$, $\mathbb{1}{u \sim v} \sim Bernoulli(p)$.

 $G \sim HSBM(V, \vec{m}, B)$, where $\vec{m} = [\vec{m}_1^\top \mid \vec{m}_2^\top \mid \cdots \mid \vec{m}_R^\top]^\top$ and B is given by

$$B = \begin{bmatrix} B_1 & pJ_{K,K} & \cdots & pJ_{K,K} \\ pJ_{K,K} & B_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & pJ_{K,K} \\ pJ_{K,K} & \cdots & pJ_{K,K} & B_R \end{bmatrix}$$

Discussion

Hypothesis

Given generating block-probability matrix $B^* \in \mathcal{B}_K$ and $\epsilon \in [0, 1]$, we let $B_r \stackrel{iid}{\sim} F_{B^*,\epsilon}$ where $F_{B^*,\epsilon}$ specifies mutually independent entries with

 $B_{r,ij} \sim Uniform(B_{ij}^* \pm \epsilon \bar{u}_{ij})$

for $i \leq j$ (recall symmetry of B_r). Here $\bar{u}_{ij} = \min\{B_{ij}^*, 1 - B_{ij}^*\}$ is the maximum allowable range guaranteeing that the random variable $B_{r,ij}$ stays in [0, 1]. Hypotheses of interest are given by

 $H_0: \epsilon \ge \epsilon_0$

versus

$$H_A: \epsilon < \epsilon_0$$

with $\varepsilon_0\in(0,1].$ A smaller value of ε corresponds to motifs having more coherent structure.

Biologically Motivated Model

Our cortical graph G:

- *n* vertices, with $n \in \{10^3, 10^4, 10^5, 10^6\}$.
- V = [n] is partitioned into R subsets $\{V_r\}_{r=1}^R$.
- $H_r = \Omega(V_r)$ each have $|V_r| = n_r = m = 100$, K = 5 blocks.
- R = n/m.

• $\vec{m}_r = [2, 50, 15, 8, 25]^\top \forall r.$

$$B^* = \begin{bmatrix} 0.1 & 0.045 & 0.015 & 0.19 & 0\\ 0.045 & 0.05 & 0.035 & 0.14 & 0.03\\ 0.015 & 0.035 & 0.08 & 0.105 & 0.04\\ 0.19 & 0.14 & 0.105 & 0.29 & 0.13\\ 0 & 0.03 & 0.04 & 0.13 & 0.09 \end{bmatrix}$$

•
$$p \in \{10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}\}.$$

Izhikevich E. M. and Edelman G. M. (2008) "Large-Scale Model of Mammalian Thalamocortical Systems," PNAS, 105:3593-3598

Large-scale model of mammalian thalamocortical systems

Eugene M. Izhikevich and Gerald M. Edelman*

The Neurosciences Institute, 10640 John Jay Hopkins Drive, San Diego, CA 92121

Contributed by Gerald M. Edelman, December 27, 2007 (sent for review December 21, 2007)

The understanding of the structural and dynamic complexity of mammalian brains is greatly facilitated by computer simulations. We present here a detailed large-scale thalamocortical model based on experimental measures in several mammalian species. The model spans three anatomical scales. (i) It is based on global (white-matter) thalamocortical anatomy obtained by means of diffusion tensor imaging (DTI) of a human brain. (ii) It includes multiple thalamic nuclei and six-lavered cortical microcircuitry based on in vitro labeling and three-dimensional reconstruction of single neurons of cat visual cortex. (iii) It has 22 basic types of neurons with appropriate laminar distribution of their branching dendritic trees. The model simulates one million multicompartmental spiking neurons calibrated to reproduce known types of responses recorded in vitro in rats. It has almost half a billion synapses with appropriate receptor kinetics, short-term plasticity, and long-term dendritic spike-timing-dependent synaptic plasticity (dendritic STDP). The model exhibits behavioral regimes of normal brain activity that were not explicitly built-in but emerged spontaneously as the result of interactions among anatomical and dynamic processes. We describe spontaneous activity, sensitivity to changes in individual neurons, emergence of waves and rhythms, and functional connectivity on different scales.

brain models | cerebral cortex | diffusion tensor imaging | oscillations | spike-timing-dependent synaptic plasticity



Fig. 1. The model's global thalamocortical geometry and white matter anatomy was obtained by means of diffusion tensor imaging (DTI) of a normal human brain. In the illustration, left frontal, parietal, and a part of temporal cortex have been cut to show a small fraction of white-matter fibers, colorcoded according to their destination.



PNAS

Izhikevich E. M. and Edelman G. M. (2008) "Large-Scale Model of Mammalian Thalamocortical Systems," PNAS, 105:3593-3598

and a start of a start when an a start all a start and a start and a start and a start and and and a start and a	4.1	o the they
* * * * * * * * * * * * * * * * * * * *	4.1	
nb1 1.5 8890 10.1 6.3 0.6 1.1 0.1 0.1		
p2/3L28 26 5800 . 59.9 9.1 4.4 0.6 6.9 7.7 . 0.8 7.4 2.3 0.8		
	4.1	
b2/3 3.1 3854 1.3 51.6 10.6 3.4 0.5 5.8 6.6 . 0.8 6.3 2.1 0.7 9.8 .	0.5 .	
nb2/3 4.2 3307 1.7 48.6 11.4 3.3 0.5 5.5 6.2 . 0.8 5.9 1.8 0.6 13 .	0.7	
ss4(L4) 9.2 5792 . 2.7 0.2 0.6 11.9 3.7 4.1 7.1 2 0.8 0.1 32.7 5.8 25.3 1.7	1.3 .	
ss4(L2/3) 9.2 4989 . 5.6 0.4 0.8 11.3 3.8 4.3 7.2 2.1 1.1 0.1 31.1 5.5 23.9 1.7	1.3 .	
p414 9.2 5031 . 4.3 0.2 0.6 11.5 3.6 4.2 7.2 2.1 1.2 0.1 31.4 0.1 . 5.9 24.5 1.7	1.3	
	1. 1.	
	4.1	
D4 5.4 3230 . 5.8 0.5 0.8 11 3.8 4.2 8.4 2.4 1.1	1.2	
ND4 1.5 3688 . 27 0.2 0.6 11.7 3.6 4 8.2 2.3 0.8 0.1 32.2 5.7 24.9 1.7	1.3	
2 p5(L2/3) 4.8 4316 . 45.9 1.8 0.3 3.3 2 7.5 . 0.9 11.7 1 0.8 1.1 2.3 2.1 . 11.5 7.2 0.1	0.4 .	
D 14 283 . 2.8 0.1 0.7 122 3.8 4.2 5.2 1.5 0.8 0.1 33.7 5.9 26 1.8	1.4 .	· ·
	4.1	
	9.1	
Delegation 1.3 Store 1.444.3 1.7 0.2 3.2 2 7.3 . 0.8 11.3 1.2 0.8 1.1 2.3 2.3 0.3 11.3 9.2 0.2	0.5 .	
	1.9	
	4.1	
	9.1	
	0.4 .	
n 105 0.6 2961 . 45.5 2.5 0.2 5.5 2 7.5 . 1.1 11.6 1 0.9 1.5 2.5 2 . 11.4 7.2 0.1 5 pc/14/14 12 6 2961 . 25 0.1 0.1 0.7 0.0 12	2.0	
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	2.9 .	
	1.4	
	1.4	
	20	
ph(L3/0) 4.3 3.3 2.5 0.1 0.1 0.7 0.5 1.3 . 0.1 0.1 4.5 . 0.3 1.2 1.2 1.2 1.0 1.0 1.0 1.1 1.2 1.2 1.2 1.2 1.1 1.2 1.2 1.2 1.2	0.4	
	1.4	
	1.4	
	41	
b6 2 3220 25 01 01 07 09 13 01 01 49 04 12 132 77 77 557 06	20	
	20	
	2.0	
TCs 0.5 4000 31 71 23 8	5	25.9
TCn 0.5 4000 31 71 1 1 1 1 38 1 1 132		5 25.9
Tis 0.1 3000 13.5 . 48.7	. 24	.4
Tin 0.1 3000 13.4 . 48.7	0.6	24.4
TRN 0.5 4000 40	10	. 10

Izhikevich E. M. and Edelman G. M. (2008) "Large-Scale Model of Mammalian Thalamocortical Systems," PNAS, 105:3593-3598



(b) organized null



Cortical Column Conjecture Test

Identifying Large Scale Structures

Cortical Column Conjecture Test

Let A_r be the adjacency matrix associated with induced subgraph H_r .

We consider the test statistic

$$T = \sum_{r=1}^{R} ||A_r - \bar{A}||.$$

(Small values of T provide evidence in favor of H_A .)

-

Cortical Column Conjecture Test $0 \bullet 00$

Identifying Large Scale Structures

Discussion

If the H_r are known and aligned



For example, the *p*-value for the alternative observation, with $T \approx 48.17$, is p = 0.00345, justifying our claim that "the repeated motif structure is evident."

Cortical Column Conjecture Test $\circ \circ \bullet \circ$

Identifying Large Scale Structures

Discussion

The Effect of Graph Matching

For the test statistic

$$T = \sum_{r=1}^{R} \|A_r - \bar{A}_r\|$$

we use post-matching sample graph means

$$\bar{A}_r = (1/R) \sum_{r'} \tilde{A}_{r,r'}.$$

Here

$$\tilde{A}_{r,r'} = \operatorname*{arg\,min}_{P \in P(n_r \lor n_{r'})} \|A_r - PA_{r'}P^\top\|_F$$

is the adjacency matrix for the graph $H_{r'}$ matched to the graph H_r . P(n) is the set of $n \times n$ permutation matrices (padding adjacency matrices with zeros as needed to make the matrix multiplication consistent).

Cortical Column Conjecture Test $\circ \circ \circ \bullet$

Identifying Large Scale Structures 00000 Discussion

The Effect of Graph Matching



Consider two cases:

- 1. *identity*: $\beta \approx 0.865$
- 2. *barycenter*: $\beta \approx 0.849$

Both are significantly superior to power without graph matching $(\beta \approx 0.739)$!?

- So far, we have considered the situation in which we know the true subgraphs {H_r = Ω(V_r)}^R_{r=1}.
- In practice, we must first identify these large-scale structures in *G*, *e.g.*, community detection.
- We consider the case where we have imperfectly identified subgraphs $\{\hat{H}_r = \Omega(\hat{V}_r)\}_{r=1}^{\hat{R}}$, and it will be the collection $\{\hat{H}_r\}$ that we ultimately must use to test the cortical column conjecture.

Cortical Column Conjecture Test

Identifying Large Scale Structures

••••••

Discussion

Robustness



Consider two cases:

- 1. blue: using true subgraphs $H_r = \Omega(V_r)$.
- 2. red: using imperfectly identified subgraphs $\{\hat{H}_r\}_{r=1}^R$.

NB: significant reduction of power!

Cortical Column Conjecture Test

Identifying Large Scale Structures

Community Detection

We now consider the effect of using community detection algorithms for (necessarily errorful) identification of large-scale structures (candidate motifs) in G.

- Louvain (Walktrap, Infomap, ...).
- clustering \circ adjacency spectral embedding.
- Adjusted Rand Index as a performance measure?
- Subsequent power as the performance measure!



V. D. Blondel, J.L. Guillaume, R. Lambiotte, and E. Lefebvre. "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008, 2008.

Cortical Column Conjecture Test

Identifying Large Scale Structures $\circ \circ \bullet \circ \circ$

Discussion

Example

```
Consider n = 10^3, R = 10, and \epsilon_{H_0} = 1.0, \epsilon_{H_A} = 0.3.
```

```
With p = 0.001:
```

```
## Louvain
wcL <- multilevel.community(g)
membL <- membership(wcL)
table(membL)
# H0:
# 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
# 1 96 1 1 1 1 92 1 103 98 105 1 94 106 98 101 100
# HA:
# 1 2 3 4 5 6 7 8 9 10 11 12
#101 1 101 101 1 95 106 96 99 100 99 100
(ordit 6 adjustedDeadTades(lebC membly))</pre>
```

(ariL <- adjustedRandIndex(labG, membL))
#[1] 0.8965267 : H0
#[1] 0.9319579 : HA</pre>

We truncate the collection of identified subgraphs $\{\hat{H}_r\}_{r=1}^{\hat{R}}$ by discarding all those "tiny" clusters, yielding $\hat{\underline{R}} \leq \hat{R}$. We proceed with the truncated collection $\{\hat{H}_r\}_{r=1}^{\hat{\underline{R}}}$. (We use $\tau = 10$.)

Cortical Column Conjecture Test

Identifying Large Scale Structures $\circ \circ \circ \bullet \circ$

Discussion

Example

algorithm	$ARI(H_0, H_A)$	power
Louvain	(0.90,0.93)	0.81
Walktrap	(0.88,0.92)	0.82
Infomap	(0.86,0.92)	0.86
skmeans \circ ase	(0.85,0.91)	0.76

Example





p = 0.001, $\beta = 0.813$







log10(p)

Identifying Large Scale Structures

Discussion

Discussion

- Ky Fan & Schatten
- estimation: $\hat{\varepsilon}$
- $(p-c)J_{K,K} + qcI_K$
- directed, weighted, loopy, multi, ...
- soft seeds: 80/20 vertex attribute split (excitatory/inhibitory)
- a limited set of computing primitives; that is, ≥ 1 repeated motifs
- extension of this method to allow for overlapping subgraphs; that is, a given node can participate in multiple instances of a repeated subgraph
- graph-centric computer vision
- graph anova
- invariants
- scalability
- QQ / errroful / experimental design

^{• ...}

Matching C.elegans



Figure 4: Plotting the matched ratio $R_m \pm 2 \, s.e.$ for matching the 253 vertex chemical and electrical C. elegans connectomes for seed values m = 0, 50, 100, 150, 200. We show the performance of the JOFC and SGM algorithms for matching the graphs for all combinations of with/without directness and with/without edge weights. JOFC is plotted in green, SGM in red, and chance in black. Note that JOFC for each combination of with/without directedness and with/without edge weights significantly outperforms the best SGM combination (directed and unweighted). For each combination of m and ρ , we ran 100 MC replicates.



Vince Lyzinski, Sancar Adali, Joshua T. Vogelstein, Youngser Park, Carey E. Priebe, "Seeded Graph Matching Via Joint Optimization of Fidelity and Commensurability," arXiv, 2014.

Cortical Column Conjecture Test

Identifying Large Scale Structures

Discussion

Multiple Repeated Motifs

After performing SGM & Louvain procedures as needed, provided that the number k of repeated motifs is \hat{k} ,

$$T = \sum_{r=1}^{R} \|A_r - \widehat{A}_r\|,$$

where $\widehat{A}_r := \sum_{i=1}^{\widehat{k}} \widehat{G}_i \widehat{H}_{ir}$, each $n \times n$ non-negative matrix \widehat{G}_i represents an (estimated) repeated motif, and each $(\widehat{H}_{1r}, \ldots, \widehat{H}_{\widehat{k}r})$ is a probability vector as estimated using the following procedures:

SVT+NMF @ rank k

Model Selection

$$\begin{aligned} Xe_r \leftarrow \operatorname{vec}(G(r)) \text{ for } r &= 1, \dots, R\\ \widehat{X} \leftarrow (U\Sigma V^\top \text{ of } X @ \operatorname{rank } k)^+\\ (\widehat{W}, \widehat{H}) \leftarrow \arg\min \|\widehat{X} - \widehat{W}\widehat{H}\|_F \end{aligned}$$

 $\widehat{k} \leftarrow \underset{k=1,...,R}{\operatorname{arg\,min}} \operatorname{AIC}(k),$ $\widehat{G}_i \leftarrow \operatorname{graph}(\widehat{W}e_i) \text{ for } i = 1, \ldots, \widehat{k},$ where e_i is a standard basis in $\mathbb{R}^{\widehat{r}}$.

Quantity/Quality Trade-Off



Figure 1. Power β and its derivative $\frac{\partial \beta}{\partial x}$ as functions of the edge tracing error rate ε for our example scenario (see text for details). (We plot $(\frac{1}{2})\frac{\partial \beta}{\partial \varepsilon}(\varepsilon)$ so that the two curves are on approximately the same scale and can productively be presented on the same plot.)



C.E. Priebe, J.T. Vogelstein and D. Bock "Optimizing the quantity/quality trade-off in connectome inference," *Communications in Statistics - Theory and Methods*, Volume 42, Issue 19, pp. 3455-3462, 2013.

Hierarchical Experimental Design



Sir Ronald A. Fisher (Photograph courtesy of Antony Barrington-Brown)

The Design of Experiments

By

Sir Ronald A. Fisher, Sc.D., F.R.S.

Hunnary Resarch Follow, Division of Mathematical Statistics, CS.IR.O., University of Adalikity Forogin Associate, United States National Academy of Sciences, and Foreign Honorary Member. American Academy of Sciences, and the Royal Duslik Academy of Sciences and Letters; Member of the Fondinol Academy McMore Letters; Member of the Fondinol Academy McMore Letters; Member of the Fondinol Academy McMore Letters; Member of Calebook Sciences and Letters; Member of Letters; Member of Neurosciences Academic Sciences and Sciences and Arbitra Baloon Professor Sciences and Letters; Member of Calebook Sciences and Letters; Member of Letters; Member of Neurosciences Academic Sciences and Sciences and Arbitra Baloon Professor Sciences and Letters; Member of Letters; Member of Neurosciences Academic Sciences and Sciences and Arbitra Baloon Professor Sciences and Letters; Member of L



HAFNER PUBLISHING COMPANY New York 1971

Cortical Column Conjecture Test

Identifying Large Scale Structures

Discussion

Leopold Kronecker to Hermann von Helmholtz: "The wealth of your practical experience with sane and interesting problems will give to mathematics a new direction and a new impetus."





Kronecker

Helmholtz