EN.580.694: Statistical Connectomics
Final Project Proposal

Indigo V. L. Rose · April 2, 2015

Clustering Algorithm Reverse Engineering of the Mouse Neocortex Connectome

Opportunity Recent advances in increasingly higher resolution axon tracing techniques have enabled the compilation of whole connectomes at an unprecedented level [1]. The Mouse Connectome Project assembled a complete cortico-cortical mesoscale connectome for the mouse cortex, allowing researchers to examine the interconnectivity of the mouse neocortex. The data gleaned from this has far-reaching implications for how the brain communicates and processes information, and is the first step to understanding human cognition from the brain?s connectomic architecture [2]. Clustering the data into subnetworks allows for quasi-independent analyses of network systems, showing specialization of areas which can confirm anatomical data or suggest new approaches for how subnetworks support behaviors.

Challenge Identifying subnetworks of the connectome is not a straight-forward task and reverse engineering the algorithm is difficult because only the raw data and the image of the final, clustered connectivity matrix is available. There are many different ways to cluster data and what information is meaningful and/or useful can change based on the clustering algorithm used. This project will examine a few common algorithms and compare them to see what the paper did and what might work better.

Action The raw data will need to be copied from Zingg, Hintiryan, et al. and processed. The clustering algorithm needs to be uncovered, as this paper didn't include any information about the clustering. Starting with kmeans, I'll be building up to more sophisticated algorithms, like spectral grouping, a Markov Cluster Algorithm, and a hierarchical clustering analysis. This data will then be compared to the clusters generated in the paper to see which one was the most likely used, and what the implications of using this specific algorithm are.

Resolution The paper identified 4 major subnetworks (Somatic Sensiomotor, Medial, Lateral, and Claustrum-Entorhinal), many composed of various smaller modules (12 in total). The results of the analysis will confirm the algorithm used for generating these subnetworks and modules. It will also provide a several alternative methods, which will each have their advantages and drawbacks.

Future Work The analysis of clustering algorithm used provides a tool to understand the assumptions made by the various algorithms which will allow more careful cognizance of what clustering of nodes means for the mouse neocortex. It's also a starting point to reexamine which clustering method is really the best when dealing with this data of this time and how to implement it in future mesoscale connectomic research.

Statistical Decision Theoretic

Sample Space $(A, P) \in \mathcal{A}_N \times \mathcal{P}_N = \mathscr{G}_N$, where Adjacency Matrices: $A = \{0,1\}^{N \times N}$ and Vertex Positions: $P = \{P_i\}$, where $P_i = (x_i, y_i, z_i) \in \mathbb{R}^3$.

Model $\{SBM_N^{k_i}(\rho, \beta) : \rho \in \Delta_{k_i} \ \beta \in (0,1)^{k_i \times k_i}\}$

Action Space $\mathcal{W}_n \times \mathcal{W}_n$, where $W \in \mathcal{Z}^{n \times n}$, $\forall$ n corresponding to three different atlases, $\{Desikan, Destrieux, DKT\}$, and $\mathcal{Z} = \{0, 1, 2, ...\}$. In words, W is a weighted adjacency matrix.

Decision Rule Class $f : \mathcal{A}_N \to \mathcal{W}_n$

Loss Function $\ell(\hat{W}_i^k, \hat{W}_j^k) = ||\hat{W}_i^k - \hat{W}_j^k||_F^2$

Risk Function

$$R = E[L] = \sum_{j}^{\mathscr{G}} L(G, G_j) \cdot p(G) \ / \ |\mathscr{G}|$$

References

1. Brian Zingg, Houri Hintiryan, Lin Gou, MonicaY. Song, Maxwell Bay, MichaelS. Bienkowski, NicholasN. Foster, Seita Yamashita, Ian Bowman, ArthurW. Toga, Hong-Wei Dong, Neural Networks of the Mouse Neocortex, Cell, Volume 156, Issue 5, 27 February 2014, Pages 1096-1111, ISSN 0092-8674, http://dx.doi.org/10.1016/j.cell.2014.02.023.

2. Sporns O, Tononi G, Ktter R (2005) The Human Connectome: A Structural Description of the Human Brain. PLoS Comput Biol 1(4): e42. doi:10.1371/journal.pcbi.0010042.