Studying Neural Networks with Neural Networks

Sveekruth Sheshagiri Pai
Roll No: MS15156

A dissertation submitted for the partial fulfilment of BS-MS dual degree in Science

Under the guidence of
Dr. Satyajit Jena

June 2020

Indian Institute of Science Education and Research Mohali
Sector - 81, SAS Nagar, Mohali 140306, Punjab, India

Certificate of Examination

This is to certify that the dissertation titled **"Studying Neural Networks with Neural Networks"** submitted by **Sveekruth Sheshagiri Pai** (Reg. No. MS15156) for the partial fulfillment of BS-MS dual degree programme of the Institute, has been examined by the thesis committee duly appointed by the Institute. The committee finds the work done by the candidate satisfactory and recommends that the report be accepted.

Dr. Shashi Bhushan Pandit Dr. Ananth Venkatesan Dr. Satyajit Jena
(Supervisor)

Dated: 13.06.2020

Declaration

The work presented in this dissertation has been carried out by me under the guidance of Dr. Satyajit Jena at the Indian Institute of Science Education and Research Mohali.

This work has not been submitted in part or in full for a degree, a diploma, or a fellowship to any other university or institute. Whenever contributions of others are involved, every effort is made to indicate this clearly, with due acknowledgement of collaborative research and discussions. This thesis is a bonafide record of original work done by me and all sources listed within have been detailed in the bibliography.

Sveekruth Sheshagiri Pai

(Candidate)

Dated: June 13, 2020

In my capacity as the supervisor of the candidate's project work, I certify that the above statements by the candidate are true to the best of my knowledge.

Dr. Satyajit Jena

(Supervisor)

Acknowledgement

The progression of this thesis has been quite serendipitous. I am grateful to Dr. Satyajit Jena for being more than a thesis supervisor to me during my final year in IISER Mohali. His openness to new ideas and encouraging demeanor are what allowed me to embark on a topic of my choice. I owe a lot of this project to his unconditional support. A complete novice in theoretical and computational neuroscience, I was in dire need of a new idea that would benefit the most from his expertise in machine learning methods. It was a fateful encounter with Dr. Venkatakrishnan Ramaswamy at the Simons Centre at NCBS last summer that united my previous knowledge of connectomics with new topics from theoretical neuroscience and deep learning (DL). I gained skills in network analysis, the foundation of this work, thanks to Prof. Somdatta Sinha's meticulous mentorship last semester.

In the end, it all came together! Overall, I have learned a great deal over the span of a year, and will continue to do so in the spirit of this endeavor. I could spend pages together thanking the many unnamed souls who have helped me reach the finish line, however, pertaining to my thesis, I have special mentions for two. I am indebted to fellow EHEP lab member, Rohit Gupta, for being the lifesaver that he is, lending me his laptop to continue work when mine died abruptly in the midst of an ongoing pandemic, and for giving me company during this emptiness. Finally, I thank my mother, Deepa Kamath, for being the epitome of her name, my light in the darkest of times, and my number one supporter.

Sveekruth Sheshagiri Pai

MS15156

IISER Mohali.

List of Figures

Contents

Abstract

Brains are complex networks of neurons that enable animals to survive and thrive in their environment. They are capable of sensory transduction, information processing, memory storage, and motor output. However, the manner by which they perform these functions is not well understood. Computational neuroscientists attempt to construct models of the brain that will eventually address these, through various approaches. A promising new approach in this era of burgeoning data is the use of artificial neural networks. In this work, the author explores the applicability of artificial neural networks to the understanding of biological neural networks by modelling a simple circuit - thermotaxis, in a simple organism - *C. elegans*.

Chapter 1

Introduction

Since the advent of multicellular life on Earth, communication between cells became vital for the survival of organisms. Cells use electrochemical signalling in order to respond to stimuli, both external and internal to the organism, through various types of membrane transport proteins. Prior to the evolution of the nervous system, long distance communication was very limited, as most of it depended on passive processes such as simple diffusion or water flow to transport ions or chemical messengers.

The basic unit of the nervous system is the neuron (Figure 1.1), a cell specialized for electrical conduction. While there exist anatomical and physiological differences between neurons within and between organisms, their defining features include the presence of long cytoplasmic processes called neurites (dendrites and axons), gated ion channels and transmembrane pumps that help maintain an electrochemical gradient, and defined contact points between neurons at axon terminals termed as synapses. Unlike conventional modes of cellular communication, these features allow the nervous system to behave like an electrical circuit, which processes signals in a rapid, coordinated manner.

Such a system offers a significant evolutionary advantage to animals that can leverage it. Organisms using a nervous system are able to respond quickly to external stimuli, usually in the context of locomotion. For example, an escape response to predators is well characterized in several species (Figure 1.2). At the simplest level, behaviors are due to a neural circuit composed of sensory neurons (that detect a stimulus), motor neurons (that synapse onto muscles, enabling movement), and interneurons that bridge and modulate the two. As

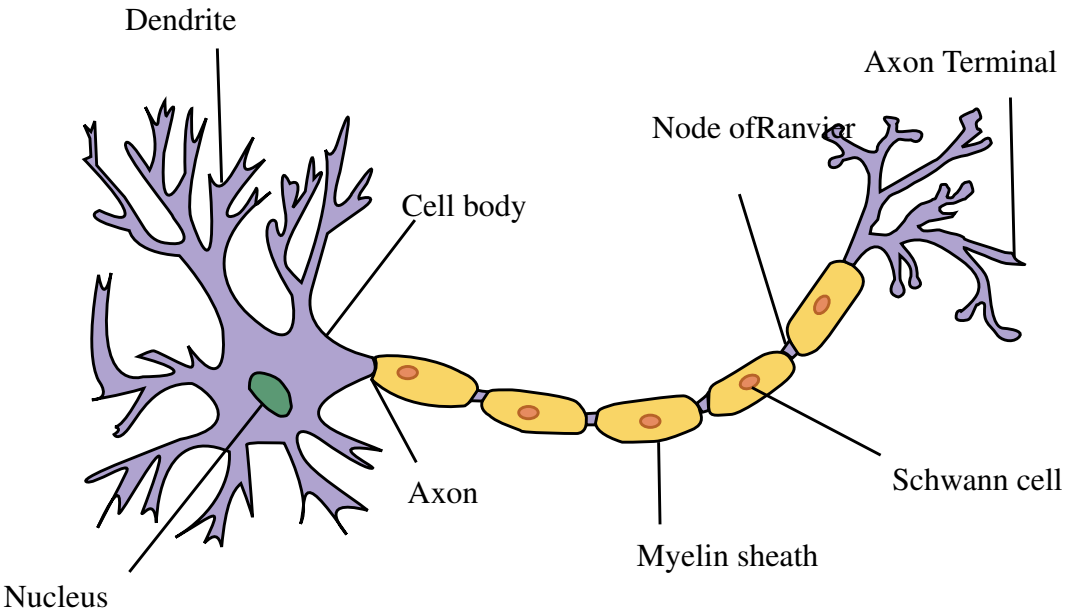


Figure 1.1: Vertebrate Multipolar Neuron [Dhp19]. Electrical signals are conducted from dendrites (left) to axon terminals (right) via the axon.

multicellular organisms grew in size, the nervous system became a persistent feature.

Nearly all extant animal species have nervous systems. Its importance is conjectured from phylogenetic relationships, suggesting an approximate ten-fold increase in neuron count every 100 million years (Figure 1.3, corresponding data in Table 1.1). This isn't enough to draw a conclusion, after all, confounds such as allometric scaling and network complexity exist. However, the trend of increasing network size across different metrics and species is worth investigating. Network science offers insight into the topology of nervous systems, and how they might have upscaled over evolutionary time.

As defined by Nature: *Computational neuroscience is the field of study in which mathematical tools and theories are used to investigate brain function. It can also incorporate diverse approaches from electrical engineering, computer science and physics in order to understand how the nervous system processes information* [noa]. It serves to complement data obtained from experimental neuroscience approaches, using a model. As with much of science, there is no unified framework till date that explains how the brain works. (Note: the word brain is often used interchangeably with the nervous system, but the former is a part of the latter, located in the head region of more recent species. Unless otherwise mentioned, the author will default to this standard practice as the term 'brain' is terse and

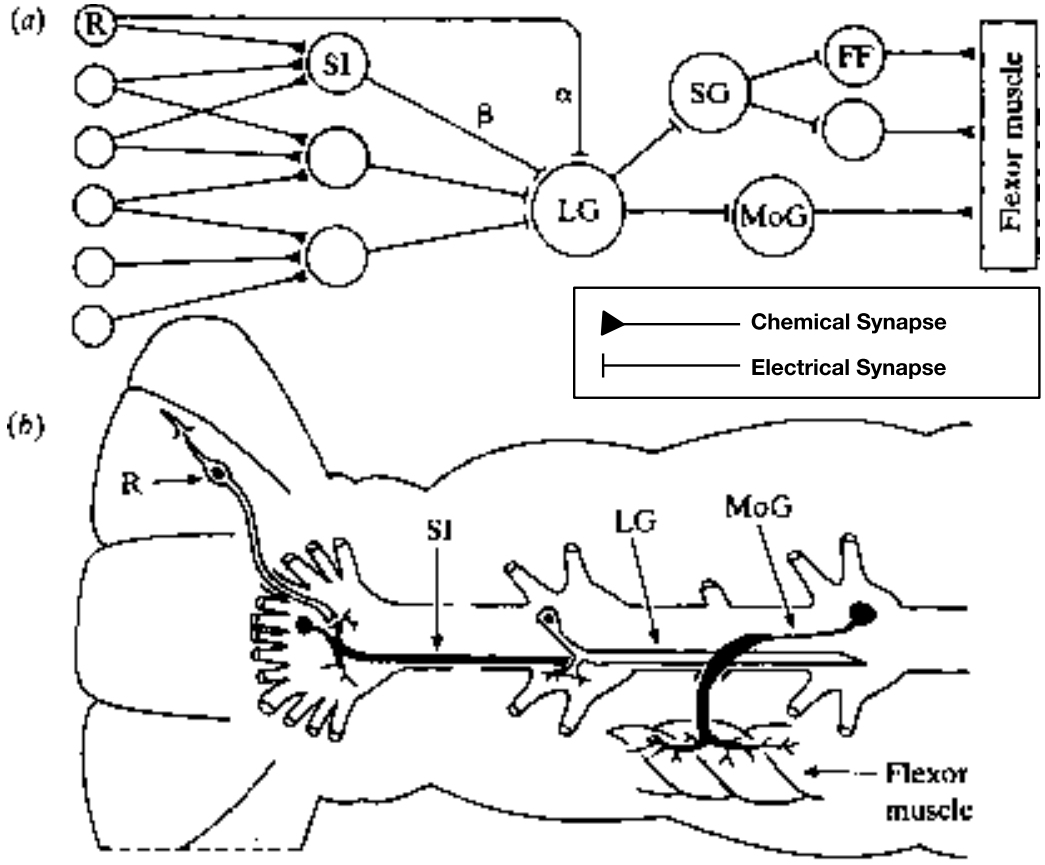


Figure 1.2: Crayfish Escape Circuit. Touch or high frequency water movements cause sensory receptors (R) to fire simultaneously, exciting downstream sensory interneurons (SI), which further excite the lateral giant (LG) and other command interneurons (Section 2.3), which are sufficient to elicit a motor response (FF - fast flexor motor neurons) known as 'tail-flipping'. (a) Schematic circuit. Chemical and electrical synapses (Section 3.2). (b) Anatomic circuit.

ubiquitous in usage).

Researchers in the mid 20th century realized that artificial neurons loosely based on their biological counterparts were capable of universal computation. Viewed as more generalized versions of their logic gate predecessors, computer scientists had already proven the universality of NAND and NOR gates in forming the basis of Boolean algebra, and consequently, all of computation [Nie15]. However, unlike logic gates, their parameters need to be set appropriately before they become useful. The ingenious work of several researchers from the 80s onward gave rise to the field of deep learning, which is essentially a branch of machine learning methods that train these parameters in multilayer artificial neural networks (ANNs) to learn patterns in data.

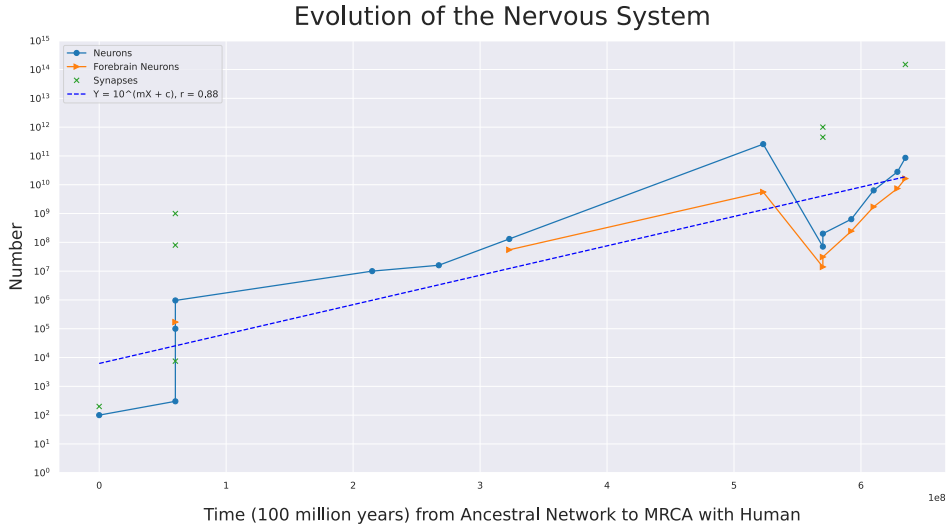


Figure 1.3: Evolution of the Nervous System. Semi Log Y plot of Number vs. Time. Time corresponds to the number of years (10^8) since the first known organisms with nervous systems (ctenophores), and the most recent common ancestor (MRCA) between an organism and humans. Number corresponds to neuron and synapse counts in that organism. Organism names, Time, and Number values taken from Table 1.1. The plot takes the form of an exponential relationship $Y = 10^{mX+C}$. A linear regression yields $m = 1.02 \cdot 10^{-8}$, which corresponds to a 10.49 fold increase in neuron count every 100 million years.

In recent years, a growing body of work has been published using ANNs to model biological neural networks (BNNs). Connectionism, as opposed to computationalism/symbolism, is the term given to this paradigm. Here, the biophysical functions of individual neurons are abstracted in exchange for a model that explains behavior as a consequence of connection weighted non-linear transformations to input. While critics of this approach emphasize that ANNs and their learning methods are biologically implausible, the author recognizes that a mechanistic understanding of how BNNs compute functions remains an unsolved challenge that could benefit from a fresh perspective [KBSK20]. Studying the brain's architecture carefully can help one build a more realistic ANN, with the hope of eventually revealing the secrets behind the complex wiring that explains how a BNN learns. Connectomics, as defined on Scholarpedia is: *The field of science dealing with the assembly, mapping and analysis of data on neural connections* [Spo10].

In this thesis, the author starts by presenting the reader with a broad overview of a sim-

TABLE 1.1	Ctenophore	C. elegans	Fruit Fly	Honey Bee	Zebrafish	Frog	Zebra Finch
Time	0.00E+00	6.00E+07	6.00E+07	6.00E+07	2.15E+08	2.68E+08	3.23E+08
Neurons	*1.00E+02*	3.02E+02	1.00E+05	9.60E+05	1.00E+07	1.60E+07	1.31E+08
Synapses	2.00E+02	7.50E+03	8.00E+07	1.00E+09	NA	NA	NA
Forebrain Neurons	NA	NA	NA	1.70E+05	NA	NA	5.50E+07

TABLE 1 (cont.)	African Elephant	House Mouse	Brown Rat	Marmoset	Rhesus Macacque	Chimp-anzee	Human
Time	5.23E+08	5.70E+08	5.70E+08	5.92E+08	6.10E+08	6.29E+08	6.35E+08
Neurons	2.57E+11	7.10E+07	2.00E+08	6.36E+08	6.38E+09	2.80E+10	8.60E+10
Synapses	NA	1.00E+12	4.48E+11	NA	NA	NA	1.50E+14
Forebrain Neurons	5.60E+09	1.40E+07	3.10E+07	2.45E+08	1.71E+09	7.40E+09	1.63E+10

Table 1.1: Neuron and synapse count data [noa20] used in Figure 1.3. Up - Time increases along columns. Down - Table 1 continued. NA - Data unavailable.

ple nervous system, the hermaphrodite *C. elegans* connectome. It is one of few organisms to have its entire nervous system mapped to the synaptic level, and has been well characterized through decades of experimentation. Concepts from network science will provide a descriptive understanding of the same. Next, the author focuses on a particular BNN within this connectome, namely, the thermotaxis circuit. After summarizing its structure and experimental evidence for its involvement in temperature navigation behavior, the author attempts to model the same with an ANN. To conclude, the author discusses the results, their implications, and future directions in this area.

Chapter 2

C. elegans and Connectomics

2.1 *Caenorhabditis elegans*

Caenorhabditis elegans is a free-living nematode found in the soil of temperate regions worldwide. At an adult length of 1mm, it is primarily a bacterial feeder. It has several desirable properties that make it a model organism for biologists. For starters, it (hermaphrodite variant, see Section 2.3) has a short life cycle (2.5 days) and a short lifespan (3 weeks in normal environmental conditions). Next, it is transparent throughout development (Figure 2.1a). Third, it has a stereotyped mosaic development pattern and exhibits eutely, i.e., has a small and constant cell count on reaching maturity (Figure 2.1b). Fourth, it has high fecundity, laying 300 eggs via self-fertilization [AHW$^+$20]. Add to these the benefits of being cost effective, easy to handle, cryopreservable, and non pathogenic, and it can be seen why Sydney Brenner, the founding father of *C. elegans* biology, chose to work with this species.

These features made it the first multicellular organism to have its entire genome (1998) and connectome (1986) sequenced. Remarkably, as of 2020, it remains the only model organism to have had both these 'omes' published, and to have a connectome description prior to a genome. In all the others, including humans, whole genomes have been published, whereas connectome mapping remains far from complete, with the fruit fly (*Drosophila melanogaster*) connectome touted to be released to the public by 2021, after surpassing mammoth challenges in electron microscopy and image reconstruction technology [XJL$^+$20]. Currently, feasible methods for connectome reconstructions in mammals are unavailable, owing to the sheer enormity of the task at hand.

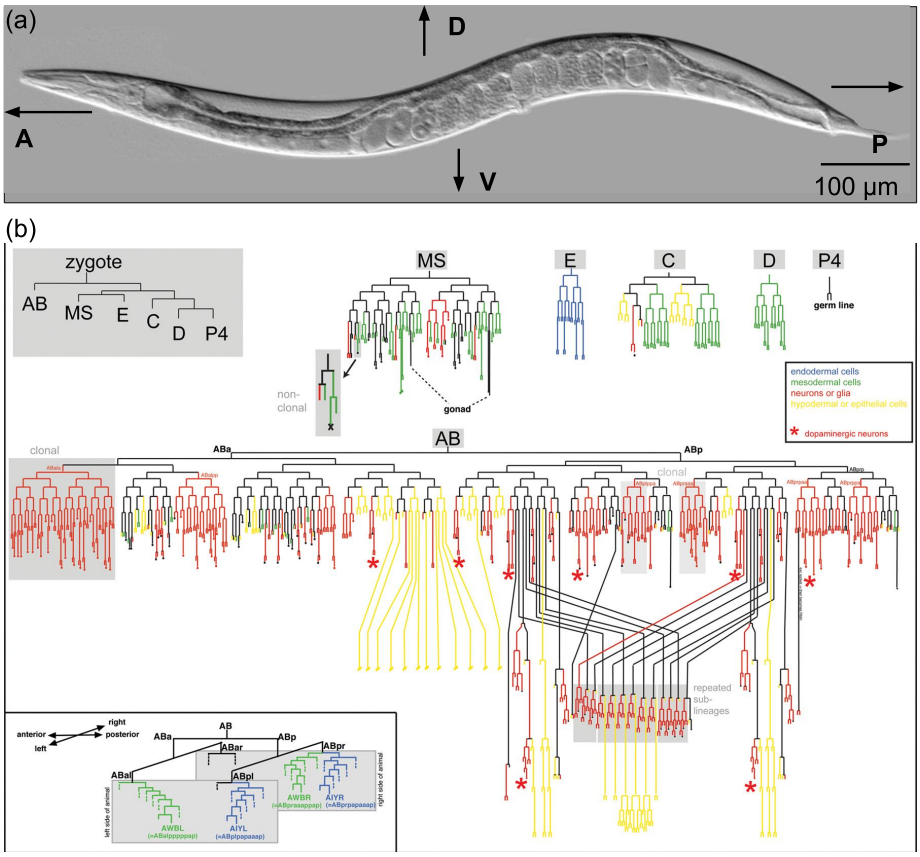


Figure 2.1: *C. elegans* Anatomy and Development.(a): Transparent hermaphrodite *C. elegans* adult worm, with anterio-posterior (A/P) and dorsoventral (D/V) axes labelled. The worm moves on one of its sides (L or R, not shown). Eggs conspicuously present on the ventral side. Adapted from [Lea]. (b): The complete lineage (1031 cells). The upper left panel indicates embryonic blastomeres. Neurons/glia are shown in red, and predominantly arise from the AB germ cell. The bottom left panel shows examples of lineage symmetry and asymmetry across the LR axis. Source [Hob10].

2.2 Connectomes

But what is a connectome? The definition of a genome is rather straightforward. It is the sequence of nitrogenous bases that makes up the genetic material of an organism. In some cases, a very long sequence, but nevertheless, one dimensional, and generated from no more than five letters. The same can't be said for a connectome. Sporns and Hagmann have defined it as *the complete description of the structural connectivity (the physical wiring) of an organism's nervous system* [Spo10]. What does this structural connectivity entail? The answer differs depending on the resolution under consideration, much like looking at the

branches of a tree. From afar, one might fail to notice anything but the trunk and its major branches. On zooming in, each branch can be seen branching further, and these branches even further, until either a limit is reached in structural definition or the resolution of the instrument being used. In the process, the trunk and its major branches are out of the field of view.

The macroscale connectome is the lowest resolution (mm - cm) of a connectome, and corresponds to white matter tracts and bundles between large anatomically segregated brain regions. DTI-MRI is typically used for this. The Human Connectome Project (HCP) is one such ongoing effort [Spo10]. The mesoscale connectome/projectome has higher resolution (μm - mm), and can distinguish circuits of neurons via axonal projections, using high resolution microscopy like STP tomography in conjunction with viral tracing. The Allen Mouse Brain Connectivity Atlas is a good example [OHN+14]. Finally, the microscale connectome has the highest resolution (nm - μm), and can capture details at the level of the synapse, with techniques for the same being limited exclusively to EM (Figure 2.2). In some sense, this is the most complete description of the connectome, both due to technological limitations in resolving characteristics of individual synapses and the fact that this resolution suffices for most ongoing modelling efforts, including the ones presented here. As seen from Fig. 2.2, the log linear relationship between resolution and information content underscores the enormity of microscale connectome datasets. Henceforth, the term connectome will imply the microscale connectome, unless mentioned otherwise. The mesoscale and macroscale connectomes are compact versions of this.

2.3 The Worm Connectome

In 1986, White et.al. published the connectome of the hermaphrodite C. elegans worm [WSTB86]. Unless specified otherwise, the author refers to this sex variant when using the convenience term 'worm'. Male worms also exist, but arise infrequently in the population, and are omitted from consideration. This connectome was the result of a 15 year long effort that involved manually analyzing thousands of 50nm thick transverse EM slices from 5 individual worms and identifying the synaptic contacts between each of their 302 neurons (Figure 2.3).

The connectome is divided into two distinct components, a pharyngeal nervous system

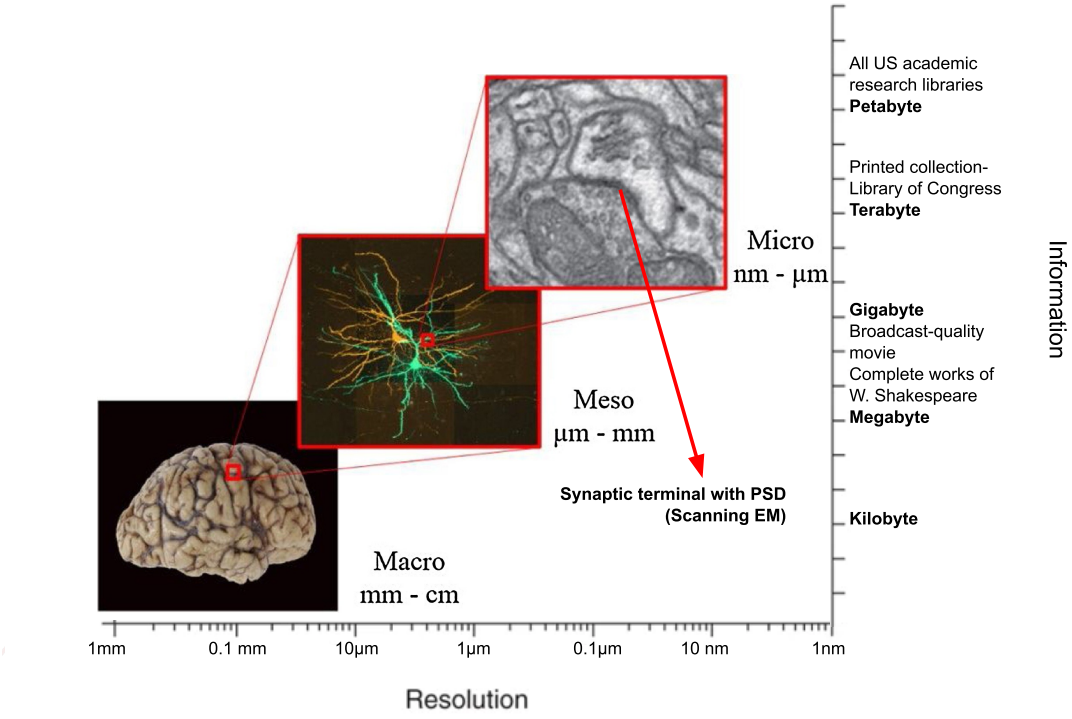


Figure 2.2: Human Connectome at Three Different Scales. A 100 fold increase in resolution results in nearly a million fold increase in storage requirements, consistent with a 3D volume. At the microscale, synaptic terminals become visible, with the postsynaptic density (PSD, a dense receptor cluster on the postsynaptic neuron) as a characteristic feature. Adapted from [BKF18] and [KL07].

with 20 neurons and a somatic nervous system 282 neurons, connected by a single pair of interneurons (RIPL/R) from the latter. We restrict our focus to the latter, which contains the neurons of the thermotaxis circuit (Section 4). This has been updated to consist of 6393 chemical synapses, 890 electrical synapses (gap junctions), and 1410 neuromuscular junctions. Of these, 3 neurons (CANL/R and VC06) are isolated. Excluding them results in a connected component of 279 neurons which is amenable to network analysis [VCP+11].

It is important to remember the motivation behind constructing the worm connectome in the first place. Brenner, an experienced geneticist, knew that gene mutations rarely had a one-to-one effect on the phenotype of an organism. This is especially true for behavior, in which the nervous system has long been implied to be the substrate. While a complete

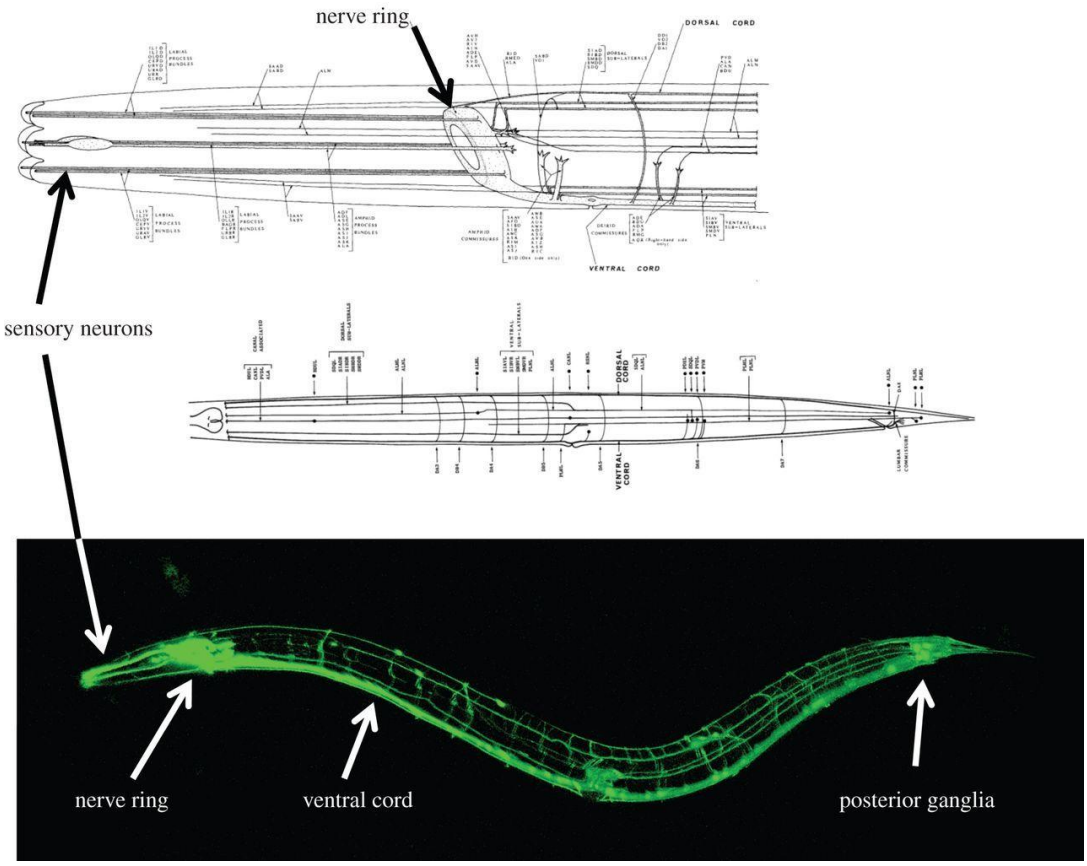


Figure 2.3: Nervous System in *C. elegans*. Top: Drawings from *The Mind of a Worm*. The head and body segments are labelled separately, showing noticeable anatomical features. Bottom: A hermaphrodite worm expressing GFP in all neurons. Most sensory neurons are located in the head and connect to the nerve ring, the 'brain' of the worm. These further connect to the ventral nerve cord which controls locomotion. Another cluster of neurons is also present in the tail, forming the posterior ganglia. Source [Emm15].

structural description is insufficient to explain functional properties in both the genome and connectome, it is necessary, as *interactions* between their parts need to be known to explain function. Hence, knowing the connectome would enable scientists to connect the dots between genes that affect behavior, via alterations to the nervous system [Emm15]. Mutants in a controlled environment would be expected to behave abnormally if the mutation affected the nervous system. Moreover, in isogenic lines, one could then posit that any changes in behavior must be due to a rewiring of the connectome from changes in the environment alone.

This line of reasoning motivated laser ablation based experiments in the worm, where

killing particular neuron classes caused quantitative changes in behavior. Albeit simple in organization, worms show a repertoire of behaviors necessary for survival, such as sensation, locomotion, and egg-laying. More complex behaviors like learning, memory, and behavior also play a role as a result. Determining a neural circuit (a subnetwork of the connectome) which is necessary for a behavior is the first step towards a mechanistic explanation for the same. Some circuits, such as the crayfish escape circuit shown earlier (Fig. 1.2), can be gleaned from inspection. This is due to the sparser connectivity of neurons in peripheral locations, and the presence of large *command neurons* [KW78], interneurons that are necessary and sufficient to elicit behavior. Others are entangled in a quagmire of connections, especially those in a cephalized location that form the brain. In the worm, the ventral motor nerve (a bundle of axons) cord that controls locomotion is an example of the former, while the nerve ring of the worm is an example of the latter.

Despite the worm being one of the smallest multicellular model organisms with a defined development plan, there is uncertainty in synapse placement across individuals (75% overlap in chemical synapses). While this makes prospects of arriving at a consensus connectome for larger organisms challenging, it highlights the role of plasticity in the brain. Changes at the synaptic level could serve as the structural substrate for learning and memory. A sufficiently large number of changes could alter connectivity patterns at the level of neurons and circuits (Figure 2.4):

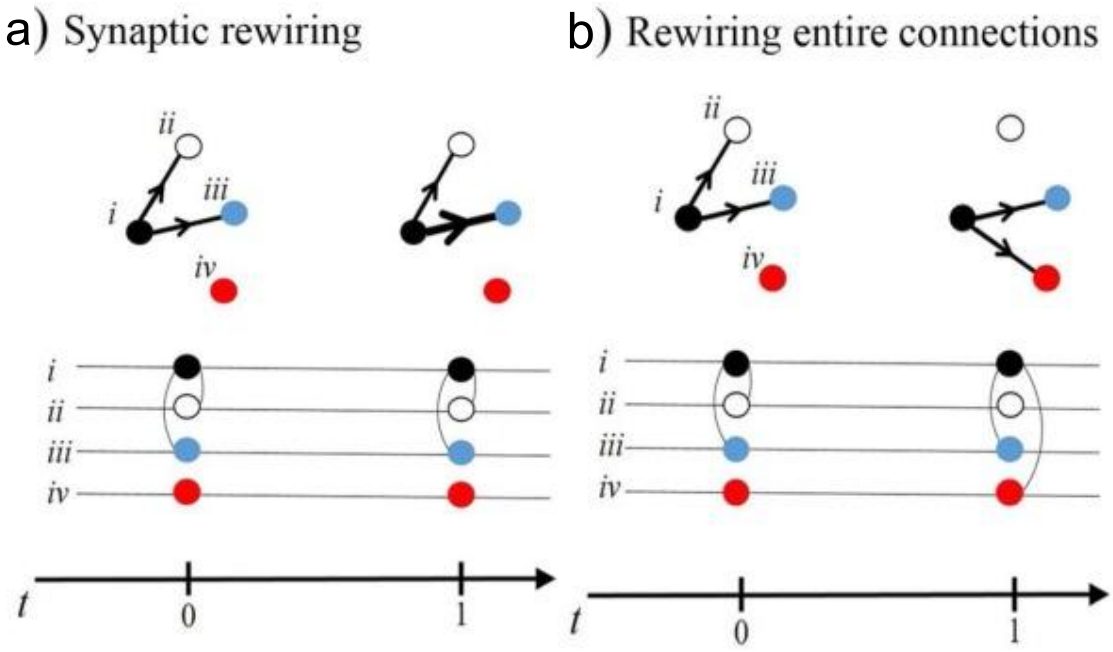


Figure 2.4: Plasticity in the Connectome, visualized by graphs (discussed in the next section), with time on the X axis. Neurons are represented as colored dots, and arrows represent synapses. (a) In synaptic plasticity/rewiring, neurons remain connected while the number of synapses between them may change. Such changes are only captured at the microscale (b) Changes in neuronal connectivity generally have more drastic implications for circuit function. Source [BKF18].

Chapter 3

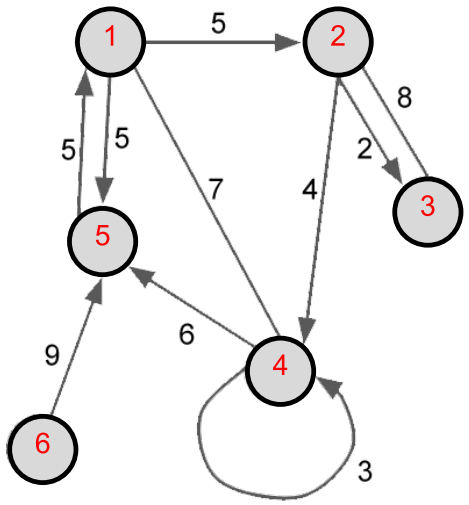
C. elegans Connectome as a Network

3.1 Networks

Network science or graph theory is the interdisciplinary study of networks. In informal terms, networks/graphs are collections of vertices/nodes and the edges/links connecting them [New10]. When each edge uniquely connects a pair of vertices, it is called a simple graph. When the order of the pair matters, there can exist two edges in each direction for a pair of vertices. Such a graph is called a simple directed graph or simple digraph. A class of graphs called mixed graphs exist, that have both types of edges. However, they add additional complexity to analyses. In such graphs, it is common practice to substitute the undirected edges for a pair of anti-parallel directed edges, without sacrificing context. Furthermore, edges can have weights, to signify a quantitative relationship between vertices. Examples include the length of a road in a road network, friendship/animosity in a social network, etc. Sometimes, it isn't enough to specify a single edge between two vertices. Multigraphs and multidigraphs are extensions to the aforementioned, that allow multiple edges to exist between vertices. Shown in Figure 3.1a is an example network of a weighted multi mixed graph, the most generalizable class.

Figure 3.1b is a convenient alternative representation of the same network, called an adjacency matrix ($\mathbf{A}$). $\mathbf{A}_{nxn}$ is a square matrix with n rows and columns, for a network with n vertices, where $\mathbf{A}_{ij}$ is an element of $\mathbf{A}$, and is the weight of the edge connecting vertex i to vertex j. While convenient and indispensable for implementing various network algorithms in this section, note that adjacency matrices fail to preserve mixed and/or multiple edge

a) Multidigraph



b) Adjacency Matrix

	1	2	3	4	5	6
1	0	5	0	7	5	0
2	0	0	10	4	0	0
3	0	8	0	0	0	0
4	7	0	0	3	6	0
5	5	0	0	0	0	0
6	0	0	0	0	9	0

Figure 3.1: Example Network. (a): Network representation, with vertices/nodes labelled from 1 - 6 (red). Grey arrows represent directed edges, while grey lines denote undirected edges. Numbers on edges (black) correspond to weights. Vertex 4 is connected to itself via a self loop. (b): Adjacency matrix representation of Fig. 3.1a. $\mathbf{A}_{23} \neq \mathbf{A}_{32}$ as a consequence of directionality and weights. Edges between vertices 1 and 5 can be replaced with a single undirected edge, conversely, the edge between 1 and 4 can be replaced with a pair of anti-parallel edges. Both vertex pairs are symmetric on $\mathbf{A}$.

information. Given $\mathbf{A}_{ij} > 1$, it is impossible to ascertain if that element represents a single weighted edge or multiple unit/weighted edges, or their directionality for that matter. As inferred from A, one solution is to convert all edges to being directed, and to sum the weights on multiple edges between a pair of vertices in a network to reduce it to a weighted simple digraph. Again, it's important to ensure that any data transformation doesn't sacrifice context, or has an appropriate explanation.

3.2 Connectome Dataset Visualized

The 279 neuron connectome dataset [VCP+11] is displayed in Figure 3.2:

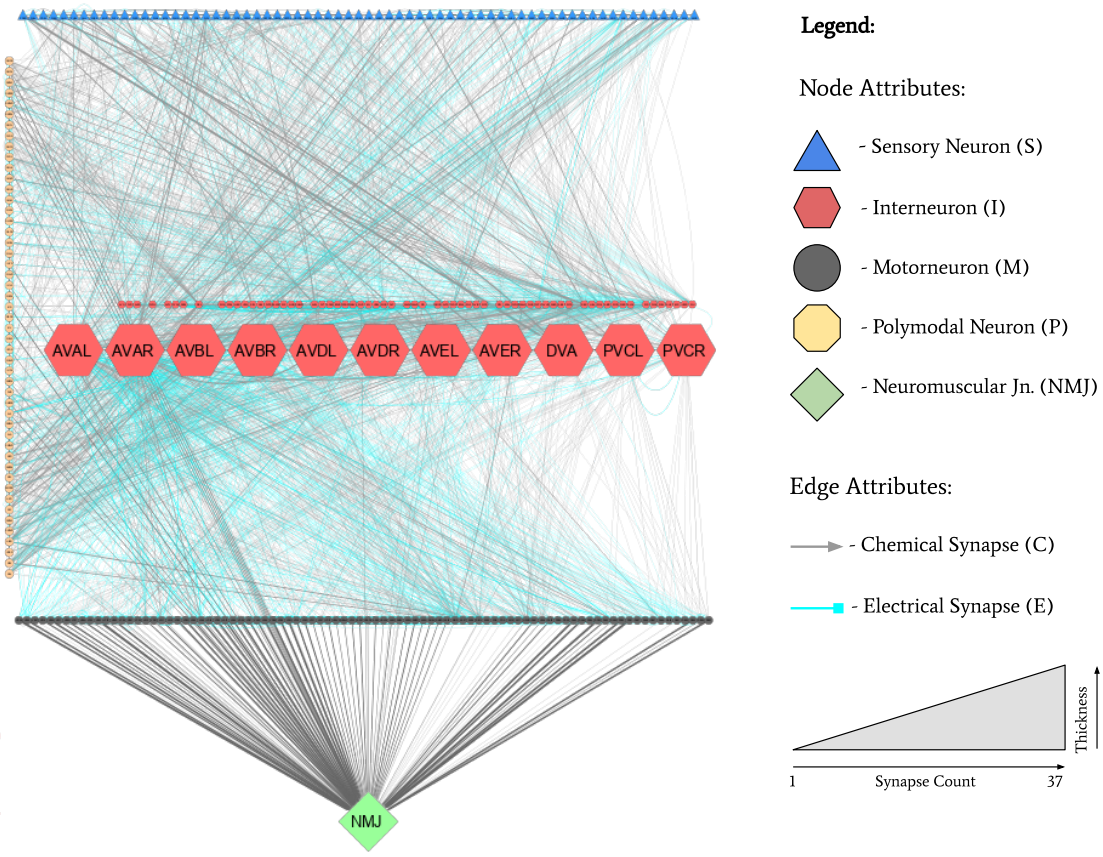


Figure 3.2: *C. elegans* Connectome [VCP+11]. Left - *C. elegans* connectome as a weighted multidigraph. Neurons are grouped in layers according to category (sensory (S), inter (I), motor (M)) roughly corresponding to signal flow hierarchy. Polymodal (P) neurons are arranged vertically on the left owing to variable placement options. The neuromuscular junction (NMJ) is located at the bottom. It, along with 11 interneurons, are shown enlarged, as they have the highest degree (discussed in the next section). Electrical (E) synapses (light blue) are distinguished from chemical (C, grey) synapses. Right - Legend, with all vertex and edge details. Prepared using [SMO+03].

In addition to sensory (S), inter (I), and motor (M) neurons, some neurons perform multiple functions and are classified as polymodal (P) [AHW+20]. The edges connecting these layers represent two types of synapses: chemical and electrical. Chemical synapses, as established by Cajal more than a century ago in what is now known as the *neuron doctrine*, are represented with directed edges representing unidirectional flows of neurotransmitters. The postsynaptic neuron's dendrite has a dense aggregation of receptors here, termed the *postsynaptic density* (PSD, refer Fig. 2.2). In contrast, electrical synapses are represented with undirected edges. They carry information bidirectionally via gap junctions, creating

TABLE 3.1	Count		Count		Count
Vertices	280	**Edges**	3339	**Synapses**	
Sensory (S)	67	Chemical (C)	2308	Chemical (C)	∼7803
Inter (I)	75	Electrical (E)	1031	Electrical (E)	∼1777
Motor (M)	88				
Polymodal (P)	49				
NMJ	1				

Table 3.1: 3.1: Neuron, edge, and synapse count from Fig. 3.2. Three electrical self loops exist. The remaining electrical edges are twice counted owing to the digraph representation. Based on data from [VCP+11]

.

reticular networks as envisioned by Golgi [Gli06]. Chemical synapses are more abundant, especially in higher organisms. They are slower than electrical synapses in signal transmission, but unlike the latter, have modulatory activity on postsynaptic neurons rather than being purely excitatory. As information is processed from S to I to M neurons, it terminates in a single vertex, the neuromuscular junction (NMJ), via chemical synapses. This vertex represents any muscle fiber triggered in this manner. The width of the edge is a proxy for the number of synaptic contacts (weight) between a pair of neurons.

In order to represent the connectome with an adjacency matrix, the author converted all electrical synapses to paired antiparallel directed edges and summed their weights with any parallel chemical synapses, essentially creating a weighted simple digraph. This is the most general version, and will be suitably converted with justifications in the analyses that follow.

3.3 Degree Centrality and Hubs

The first goal is to identify the most important vertices, and consequently, neurons in the connectome. Being a directed network with the aim of communicating electrochemical signals, a first measurement to make would be to check how many connections exist between a single neuron and its partners. This leads to the notion of *degree centrality (k)*, which is

simply the degree of a vertex. For a directed graph, a vertex has an *in-degree (k^{in})*:

$$k_j^{in} = \sum_{i=1}^{n} \mathbf{A}_{ij} \tag{3.1}$$

And an *out-degree (k^{out})*:

$$k_i^{out} = \sum_{j=1}^{n} \mathbf{A}_{ij} \tag{3.2}$$

The reader familiar with elementary linear algebra will recognize that these equations correspond to the summation of elements along a column vector $\mathbf{a}_j$ and row vector $\mathbf{a}_i$ respectively. Taking means of either n of these vectors yields the *mean degree (c)* of the network:

$$c = \frac{\sum_{j=1}^{n} k_j^{in}}{n} = \frac{\sum_{i=1}^{n} k_i^{out}}{n} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} \mathbf{A}_{ij}}{n^2} \tag{3.3}$$

The synaptic weights are inconsequential in this calculation, as the strength of coupling between a pair of neurons is not considered. Hence, all nonzero weights in **A** are set to 1, to yield an unweighted simple digraph. Plotting degree histograms and a bivariate scatterplot of k^{in} and k^{out} yields Fig. 3.3.

The histograms are both right-skewed, with mean k^{in}, k^{out} = c = 11.10. The long right tail is indicative of a power law distribution, suggestive of the Pareto principle, that few neurons in the connectome have a much larger share of connections than the majority. The networks literature term for such vertices is hubs. Moreover, the variables have a correlation coefficient of r = 0.57, which is moderately strong, suggesting that a neuron with high in-degree is likely to have a high out-degree, and vice-versa. A notable exception to this observation can be seen in NMJ, which has the highest k^{in} = 114, but the lowest k^{out} = 0. This is to be expected, as NMJ is a terminal vertex in this directed network. AVEL/R are bilateral command interneurons. Together with AVAL/R (the highest ranking hubs in both metrics) and AVDL/R, they form major drivers of the backwards locomotory circuit. Unlike the latter two however, these innervate the ventral motor nerve only till the anterior region, explaining their low total degree (Table 3.2). DVA is another interesting neuron, but with a relatively low k^{in} = 22 compared to k^{out} = 37. This is in part due to its posterior location, far from the nerve ring with a denser neuron concentration. DVA has polymodal capabilities, providing sensorimotor integration for anterior and posterior touch circuits by virtue of being stretch sensitive. This may put it higher in the directional processing hierarchy, resulting in fewer pre-synaptic partners than other interneurons with comparable out-degree.

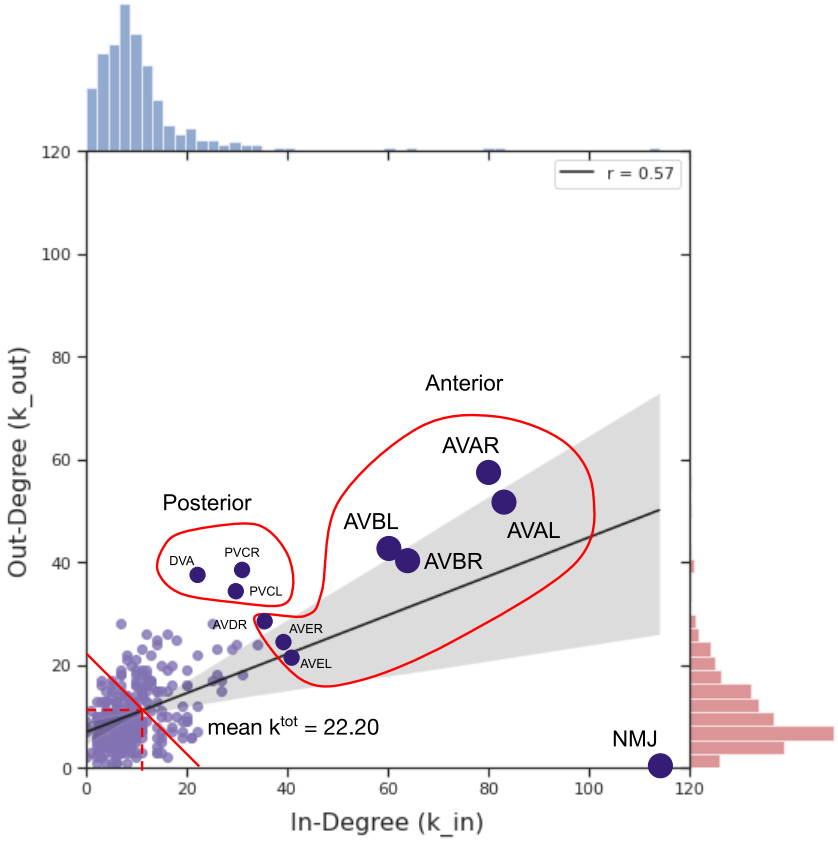


Figure 3.3: Joint Plot of Connectome in-degree (k^{in}) and out-degree (k^{out}. Most neurons (purple) occupy the bottom left corner of the plot, with low k^{in} and k^{out}. Few neurons extend further diagonally into the plot. A moderately positive correlation of r = 0.57 exists between the metrics. The top 10 neurons by total degree (see Table 3.2 below) are termed hubs (dark purple), and are located anteriorly or posteriorly. Some hubs lie on the periphery of the scatter, most notably NMJ, which has a zero out-degree owing to its terminal position in the network. Top - k^{in} histogram. Right - k^{out} histogram.

We now explore the total degree (k):

$$k = k^{in} + k^{out} \tag{3.4}$$

which is simply the sum of the in-degree and out-degree of a vertex. It is a trivial exercise to show that the mean of k for a directed network is 2c. Shown in Figure 3.4 is the resulting histogram.

As one might expect, adding two right skew distributions with a moderately positive correlation yields another right skew distribution. Here, k_{mean} = 2c = 22.20 as expected. Table 3.2 summarizes these results for the top 10 neurons and NMJ (which we remind the reader,

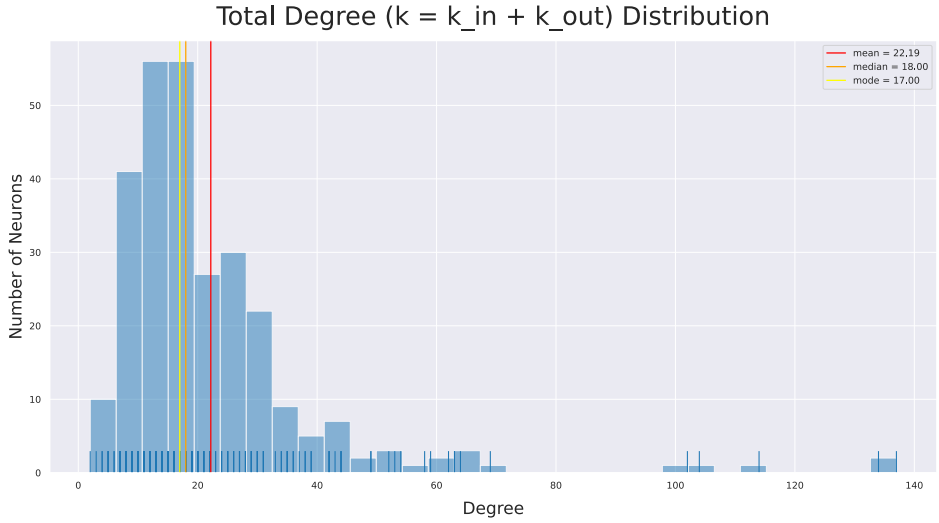


Figure 3.4: Total degree (k) distribution. As shown, k_{mean} = 22.20, and falls to the right of the peak (mode, yellow) and median (orange), making it right skewed.

isn't a neuron):

TABLE 3.2	Neuron	k	k_in	k_out
1	AVAR	**137**	80	**57**
2	AVAL	134	83	51
3	NMJ	114	114	0
4	AVBR	104	64	40
5	AVBL	102	60	42
6	PVCR	69	31	38
7	PVCL	64	30	34
8	AVER	63	39	24
9	AVDR	63	35	28
10	AVEL	62	41	21
11	DVA	59	22	37

Table 3.2: Top 11 vertices by total degree (k)

The AVB and PVC neurons are command interneurons for the forward locomotory circuit. The corresponding rug plot (1D version of a scatter plot) seems to indicate a drop in hub density after k = 40.

3.4 The Power Law and Scale-freeness

To verify the nature of the distribution in this tail region, the author employs a survival function [VCP+11]:

$$P(d) = \sum_{k=d}^{\infty} p(k) \tag{3.5}$$

Which is a complement of the cumulative distribution function. *P(d)* represents the fraction of vertices with k $\geq$, where d goes from 0 to k^{max}. Figure 3.5 shows the same, with axes log-log scaled (base 10):

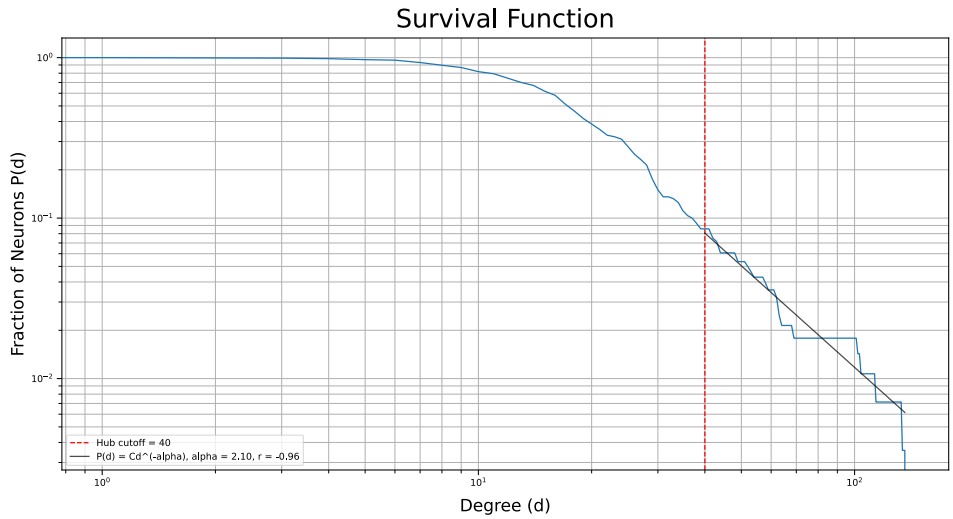


Figure 3.5: Survival function (P(d)) of total degree (k = d). Note the log scaling, vertical line at d = 40 suggesting the hub cutoff value, and linear fit (r = -0.96) with negative slope (α = 2.10), indicating a power law.

As seen in the figure, the survival function shows a linear drop-off soon after d = 10. The author employs the power law to fit the same:

$$P(d) = Cd^{-\alpha}, Range = [0, 1] \tag{3.6}$$

Where C and α are constant terms. This function has the nice property of scale invariance, i.e., P(ad) is always proportional to P(d), for a constant a. Furthermore, a log transformation yields:

$$logP(d) = -\alpha log(d) + c \tag{3.7}$$

Which is a linear function with slope -α. When applied to Figure 3.5 for d geq 40, the author found α = 2.10, with an r = 0.96 making it a good linear fit. Networks with degree

distributions obeying the power law for their high degree vertices are said to be *scale-free networks*, with 2 ¡ α ¡ 3 typically. Intriguingly, many artificial and biological networks are scale-free, such as the internet and protein-protein interaction networks. Studies with percolation theory show that scale-free networks have remarkable *fault tolerance*, i.e., random vertex failures are unlikely to impact hubs (owing to their rarity), hence the network remains connected even under those circumstances [CEbAH00]. This is important for connectomes, as this implies that the underlying structure is robust to minor random damage.

3.5 Small World Properties

Next, we turn our attention towards looking at properties concerning vertex pairs or edges. For the theory relevant to this section, we will relax this digraph to a graph. Doing so will result in several chemically synapsed neuron pairs to become bidirectional if they weren't already, however it won't affect the final observation. The geodesic distance or shortest path (d_{ij}) between two vertices i and j is the minimum number of edges that need to be traversed to get from vertex i to vertex j. The vertex mean shortest path (l_i) is given by the mean of d_{ij}:

$$l_i = \frac{\sum_{j=1}^{n} d_{ij}}{n} \tag{3.8}$$

over all vertices in the same component, which is the case with our undirected graph. Furthermore, averaging over all vertices yields the *network mean shortest path (L)*:

$$L = \frac{\sum_{i=1}^{n} l_i}{n} \tag{3.9}$$

For the undirected connectome, we get L = 2.31. The *clustering coefficient (C_i)* of a vertex i is defined as:

$$C_i = \frac{2E(N_i)}{k_i(k_i - 1)}, Range = [0, 1] \tag{3.10}$$

where E(N_i) is the number of edges between neighbors of vertex i. Since the maximum number of edges between N_i in an undirected graph is $\binom{N_i}{k}$, and N_i = k_i as the degree of i, C_i represents the fraction of connectivity between the neighbors of a node. As we have done with previous metrics, we can define the mean clustering coefficient (C):

$$C = \frac{\sum_{i=1}^{n} C_i}{n} \tag{3.11}$$

This is the Watts-Strogatz definition of C. Since C_i is inversely proportional to the square of the degree k_i, it is biased in favor of low degree nodes. For the undirected connectome,

C = 0.35. L and C are parameters that define small world networks, an important class of networks, that are defined to have:

$$L \propto log(N) \tag{3.12}$$

where N is the total number of vertices/nodes in the network. Watts and Strogatz (WS) identified them to be a category of random graphs, that could be constructed starting from an undirected radially symmetric regular graph having an even and uniform integer degree k (Figure 3.6) [WS98]:

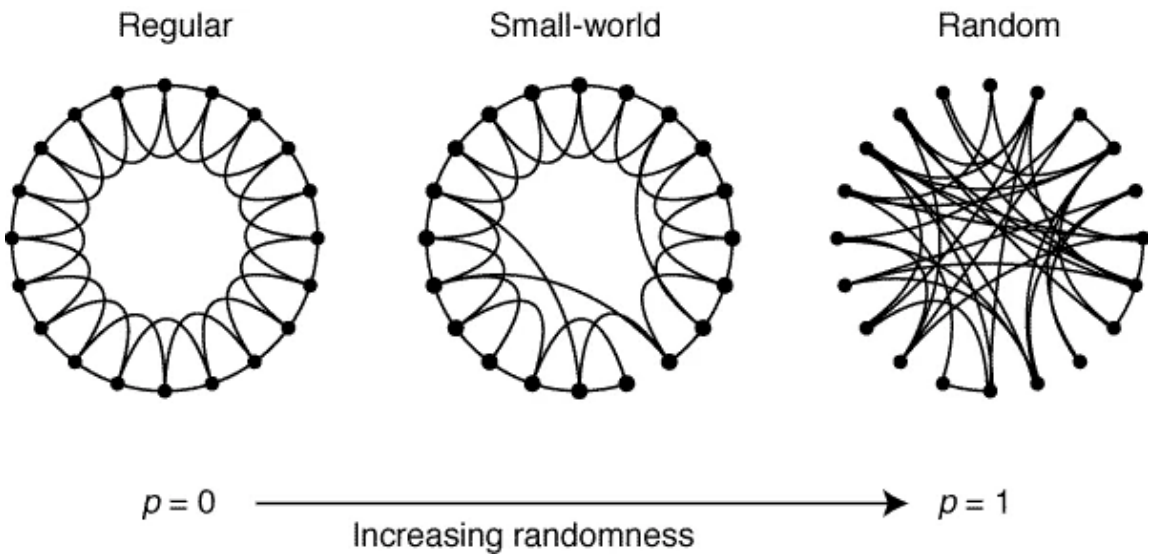


Figure 3.6: The Watts-Strogatz (WS) Model. WS model on 20 vertices starting with degree 4 each. Left - A regular graph with reassignment probability p = 0. Middle - A small-world graph with $0 < p < 1$. Right - A random graph with p = 1. Source [WS98].

Then, each edge in the network is randomly connected to any other vertex with a reassignment probability $0 \leq p \leq 1$. As seen from the figure, setting p = 0 ensures the regular graph remains, while p = 1 results in a random graph but with a fixed number of edges. From the very construction, it becomes apparent that small world networks seem to have dense local clustering, along with a handful of long range connections, making them distinct from their

random and regular counterparts which lack them respectively. This difference is described quantitatively with the *small world coefficient (S)* [HG08]:

$$S = \frac{C}{C_r} / \frac{L}{L_r} \tag{3.13}$$

Where C_r and L_r are network clustering and network mean shortest path respectively, for an equivalent random graph (p = 1). Hence, an S $\approx$ 1 indicates a random graph. To check the small worldness of the *C. elegans* connectome, we must prepare an equivalent random graph, with N = 279 and E = 2404. This yields k = 17.17, which is neither even nor an integer. Hence, the author constructed a WS random graph with k = 18, then removed edges at random until the number of edges was 2404. This process was repeated (Monte Carlo N = 100) to arrive at C_r and L_r. The same was done for a WS regular graph for comparison (while such graphs are fixed, the random removal of edges isn't, thus resulting in minor fluctuations in C_{reg} and L_{reg} between realizations). The resulting L and C values by node are displayed as density plots (Figure 3.7 and Table 3.3), which use representative realizations for the regular graph and random graph.

TABLE 3.3	**Network**	**L**	**C**	**S**
1	Random	2.25	0.06	1.00
2	*C. elegans*	2.31	0.35	5.75
3	Regular	8.27	0.67	3.12

Table 3.3: L, C, and S (small-world coefficient) for the 3 networks in Fig. 3.6

Immediately, two separate clusters of points are noticed. The regular graph vertices (green) have high L and C, while the random graph vertices (blue) have low L and C. Both histograms have a smaller spread, compared to that of the worm connectome (orange). The mean values of L_i nearly overlap between the connectome and random graphs. The major distinction between these two occurs due to C_i, which has a large spread and C lying between the expected C values of the random and regular graphs. As a result, we get S = 5.75 $>$ Sreg = 3.12 $>$ Sr = 1.00. Indeed, the worm connectome is a small world network.

One important consequence of small world networks, as their name suggests, is the fact that L scales slowly with N, implying that vertices in a network that may be separated by large physical distances are still topologically close due to long range connections. To ensure that this happens requires hubs (discussed in the next section), giving these networks

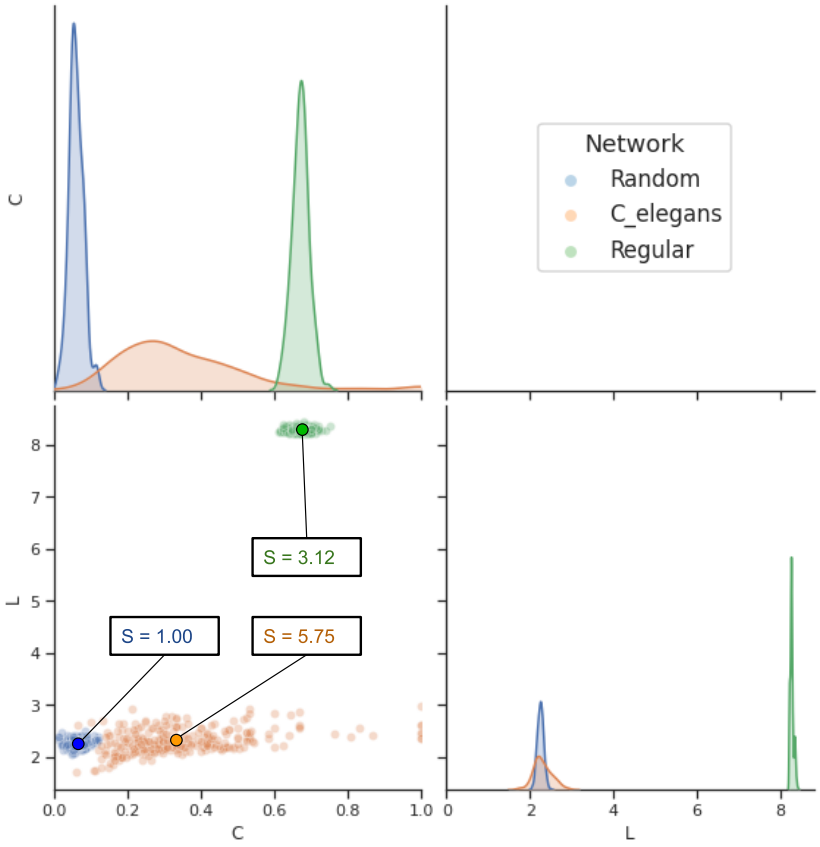


Figure 3.7: L vs. C Plot. Pair plots for mean shortest path (L) and mean clustering coefficient (C) for the *C. elegans* connectome and equivalent random and regular graphs. Bottom Left - L vs. C scatterplot. Diagonal - Kernel density histograms for L and C.

a right tailed degree distribution. Indeed, the scale-free networks mentioned earlier are special instances of small world networks that have a power law distribution, hence the undirected graph treatment applied by the author holds [CH03]. Another important consequence is *cliquishness*, or the tendency for vertices to locally organize in a compact manner with high clustering, and clique (a group of vertices that are fully connected) formation. Calculations have shown an overabundance of 2 and 3 degree cliques in the worm connectome compared to equivalent random graphs, in line with the high clustering property of small world networks [VCP+11]. This close knit structure hints at a functional role played by these motifs.

3.6 3.6 Hubs form a Rich Club

The *rich club coefficient (Φ(k))*:

$$\Phi(k) = \frac{2E_{>k}}{N_{>k}(N_{>k} - 1)} \tag{3.14}$$

is reminiscent of the clustering coefficient discussed in the previous section. It is also a measure of connectedness, but as the ratio of the number of existing edges ($E_{>k}$) to the maximum number of edges between nodes with a degree greater than k ($N_{>k}$) [MdFCC07]. At k = 0, this yields the *connectance* or *density* of an undirected graph. A plot of Φ(k) vs. k (Figure 3.8) shows a peak value of Φ(k = 44) = 0.73:

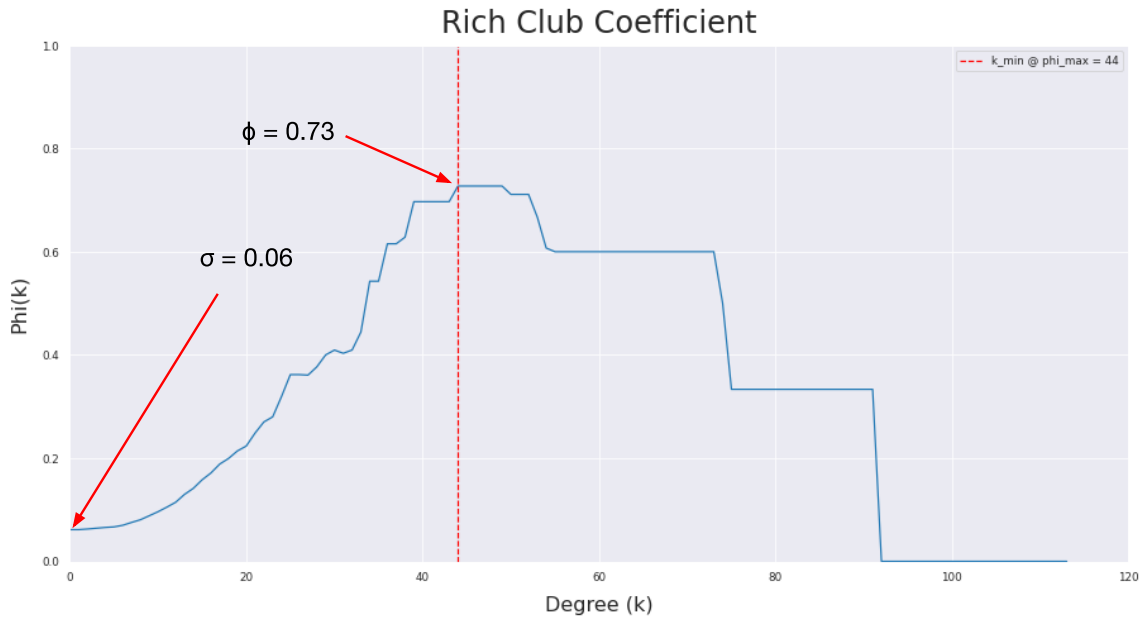


Figure 3.8: Rich Club Coefficient (Φ(k)) Plot. The minimum k (44) for Φ_{max} (0.73) is shown. Table 3.4 (below) corroborates this value as serving as a cutoff for hubs. Φ(0) = 0.06 is the connectome density. Therefore, hubs form a highly connected 'rich club'.

Doing so yields the hubs to be the same top 10 neurons from Table 3.2 ordered differently and with different values for degree owing to the undirected graph conversion applied. (AVDL having k = 44 is included in Table 3.4 as it touches the threshold). As discussed

TABLE 3.4	1	2	3	4	5	6	7	8	9	10	11	12
Neuron	NMJ	AVAR	AVAL	AVBL	AVBR	AVER	AVDR	AVEL	PVCL	PVCR	DVA	AVDL
k	114	93	92	75	74	56	55	55	54	53	50	44

Table 3.4: Neurons arranged in descending order by undirected degree ($k \geq 44$) based on cutoff from Fig. 3.8. With the exception of AVDL, all other neurons are hubs from Table 3.2.

in Section 3.3, these hubs are largely involved in locomotion, and are anterio-posteriorly (A/P) (Figure 3.9) located. This adds evidence to the long range integration provided by hubs in the connectome. Moreover, the hubs are densely connected (for comparison, the connectome density $\Phi(0) = 0.06$).

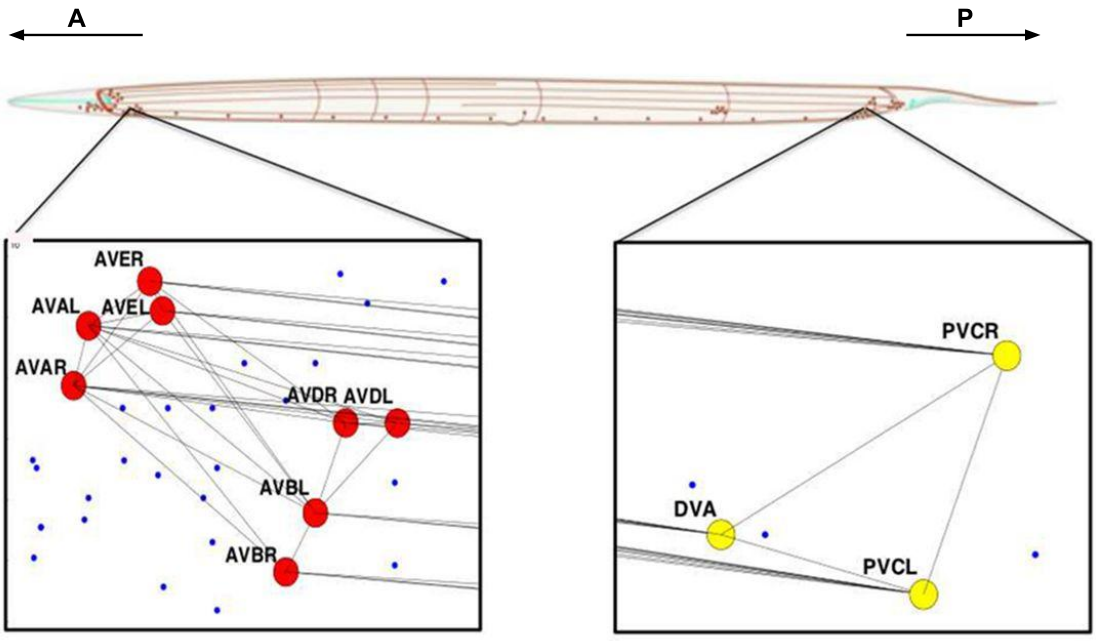


Figure 3.9: The *C. elegans* Rich Club. All hubs are concentrated anterio-posteriorly (A/P), necessitating long range connections. Small blue dots correspond to other unlabelled neurons. Adapted from [TVA+13].

3.7 Modular Organization

Finally, we look at the *modularity (Q)*:

$$Q = \frac{1}{2m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta_{c_i c_j} \tag{3.15}$$

where *m* is the number of edges in an undirected graph, δ is the Kronecker delta, and *c* denotes the category of the vertex. Breaking down the expression into its component parts, the modularity is a measurement of the fraction of total edges between two vertices minus the fraction of edges existing between them by chance provided they have the *same attributes*, i.e., they belong to the same category. Even in a network with perfect modularity, i.e., only intra attribute connections, $|Q| < 1$. Hence, a normalized version called the *attribute assortativity coefficient (r)* is preferred by the author. Here, r = 1 implies perfect assortativity, r = -1 implies perfect dissortativity (vertices have only inter attribute connections), and r = 0 is expected when connections are made at random.

Modularity is useful, because it usually underlies specific relationships between vertices. A class of algorithms in network science revolves around the problem of *community detection*, where one searches for the ideal partitioning of vertices in the network, by maximizing Q. One such algorithm is the Clauset-Newman-Moore *greedy algorithm* [CNM04], which tackles this problem by pairing all N vertices into communities over $\sim log_2 N$ steps, with each step taking $\sim N \log_2 N$ comparisons between the potential communities to choose the grouping that maximizes (or least decreases) Q at that step, and settles for the number of communities for which Q reached a maximum. Greedy optimization based algorithms aren't always guaranteed to provide optimal solutions, but make up for this with their fast run times. For a sparse graph (one with low density or having a sparse adjacency matrix) such as the worm connectome, this algorithm has a run-time of $\mathcal{O}(n \log^2 n)$, making it significantly faster than other approaches. The exponential scaling of connectomes shown earlier (Fig. 2.2) serves to remind the reader the importance of developing tractable algorithms for studying connectomes [Ram19]. Figure 3.10 shows the results of applying the greedy algorithm to the undirected connectome.

Neurons are arranged in three circular communities. With r = 0.61, this split shows moderate assortative mixing. Inter (red) and motor (black) neurons are the most abundant while polymodal (yellow) and sensory (blue) neurons are lesser in number. The smallest commu-

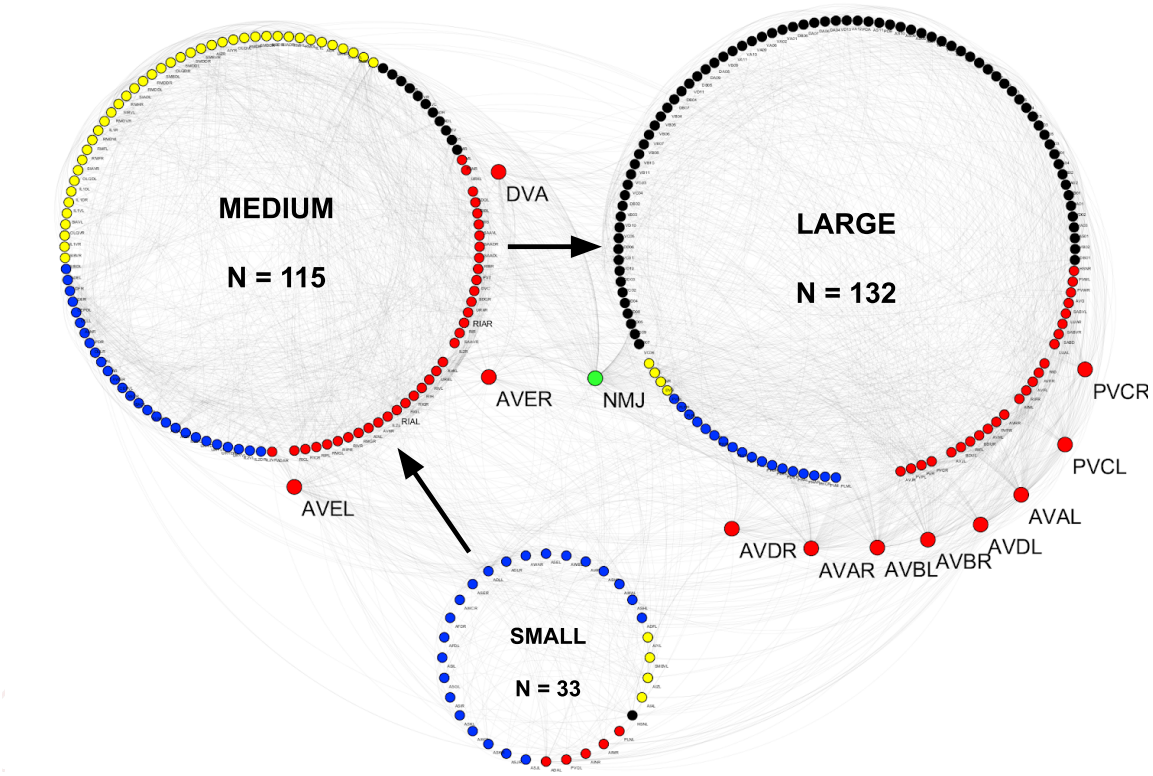


Figure 3.10: *C. elegans* Connectome Communities. Detected by greedy modularity maximization. Attribute assortativity coefficient r = 0.61. Labelling repeated from Fig. 3.2. Upper left - Medium community, having the largest fraction of polymodal neurons. Middle - Small community, dominated by amphid (Section 4.1) sensory neurons of the thermotaxis and chemotaxis circuits. Upper right - Large community, dominated by motor neurons. Arrows between communities show hypothetical signal flow.

nity has an abundance of sensory neurons responsible for thermosensation and chemosensation. The medium community has a roughly even split amongst all neuron classes (but possesses the largest fraction of polymodal neurons) with the exception of motor neurons, which become a defining feature for the largest community. This is in part due to the presence of the NMJ in this community. On looking for thermotaxis circuit (discussed in the next section) neurons, the sensory neurons including the crucial AFD were all found in the smallest community. The interneurons were divided however, in an L/R independent manner between the small and medium communities. The command neurons AVA and AVB are known to be hubs, and belong in the large community. Curiously, the hubs are also distributed in a manner that seems to be independent of anatomical axes. These observa-

tions suggest that the communities indicate a rough signal flow hierarchy, from smallest community to largest, without any obvious structural organization.

In summary, we find that the 280 neuron *C. elegans connectome* forms a sparsely connected component with small world and scale-free properties. Furthermore, all hub neurons are interneurons that integrate information over a long range, and form a rich club. These structural properties have biological underpinnings. Cross connectome comparisons have suggested that connectomes in general are sparse owing to constraints on wiring cost, but by remaining small world and scale-free, have transcended these limitations. They have modular organization with anatomical compartments formed by local neural circuits being integrated by hubs to perform complex behaviors [vdHBS16].

Chapter 4

Thermotaxis in *C. elegans*

4.1 Thermotaxis as a Vital Adaptation

Worms explore their surrounding environment via sensory neurons, mostly organized into ciliated clusters termed *sensilla* (Figure 4.1a). They help the worm to navigate gradients of temperature (thermotaxis), odor (chemo/odorsensation), and oxygen concentration (aerotaxis). Locomotory behavior in the worm is manifested as a *biased random walk* composed of variable speed *runs*, and *turns* (Figure 4.1b). The default movement of the worm is a forward sinusoid, however, curves/pirouettes ensure that they eventually approach attractive cues. Aversive/noxious stimuli result in a stereotyped pirouette with a 180° *omega turn*. The run-time to turn ratio is a function of the gradient being navigated, and increases with time if a favorable location exists nearby [GHB05]. The circuit mechanisms generating this partly stochastic behavior are what researchers are trying to elucidate. An in-depth understanding of a neuron's function in a circuit can be examined with electrophysiological techniques. The diminutive nature of *C. elegans* neurons requires special methods such as Ca^{2+} imaging and optogenetics to do so.

The *C. elegans* thermotaxis circuit is a highly studied BNN. The worm adapts to a temperature range of ($\sim$12°C-27°C)Temperatures outside this range yield noxious avoidance responses, and prolonged exposure to a heat/cold shock can result in the temporary cessation of egg-laying or even death. Within this range, worms show remarkable adaptability. When grown at a *cultivation temperature (T_C)* with adequate food, adult worms show a preference for T_C and migrate towards it. However, this preference behavior is reset to a

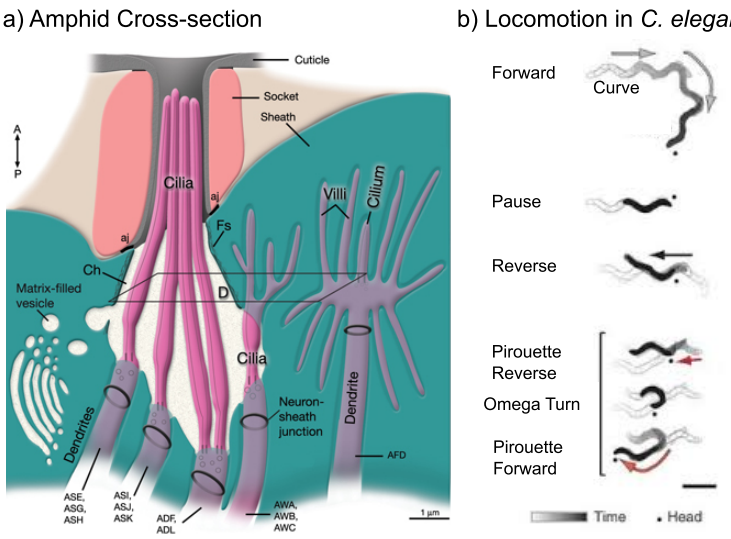


Figure 4.1: Sensorimotor Description of *C. elegans*(a) Cross section of an amphid sensillum, containing chemosensory and thermosensory (AFD, AWC) neurons. Source [AHW+20]. Ciliated endings are the site of transduction. Thermosensory neurons are embedded within the sheath cell. (b) Stereotyped movement patterns in *C. elegans*. Source [Xu19].

different temperature on starvation for 2 - 4 hours, indicating a learned association between temperature and food. This behavioral adaptation is vital to the worm's survival in its natural environment, soil, where temperature fluctuations are common. Figure 4.2 shows the typical results of thermotaxis assays.

When well fed adult worms are placed on agar plates in a linear thermal gradient as shown above, they exhibit one of three movements depending on T_C and T of their initial location: negative thermotaxis (cryophily) when $T > T_C$, positive thermotaxis (thermophily) when $T < T_C$, and isothermal tracking when $T \sim T_C$. Most thermotaxis assays quantify this behavior with the help of a thermotaxis index (TI), which has multiple definitions, but is measured on a scale of -1 to 1, and is accurate at a population level [Goo14].

4.2 The Two-Drive Model

Cell ablation studies have identified the critical componentry on the thermotaxis circuit to consist of 5 neurons (Figure 4.3).

The most important of these neurons is AFD, the primary thermosensor. Also called a *finger*

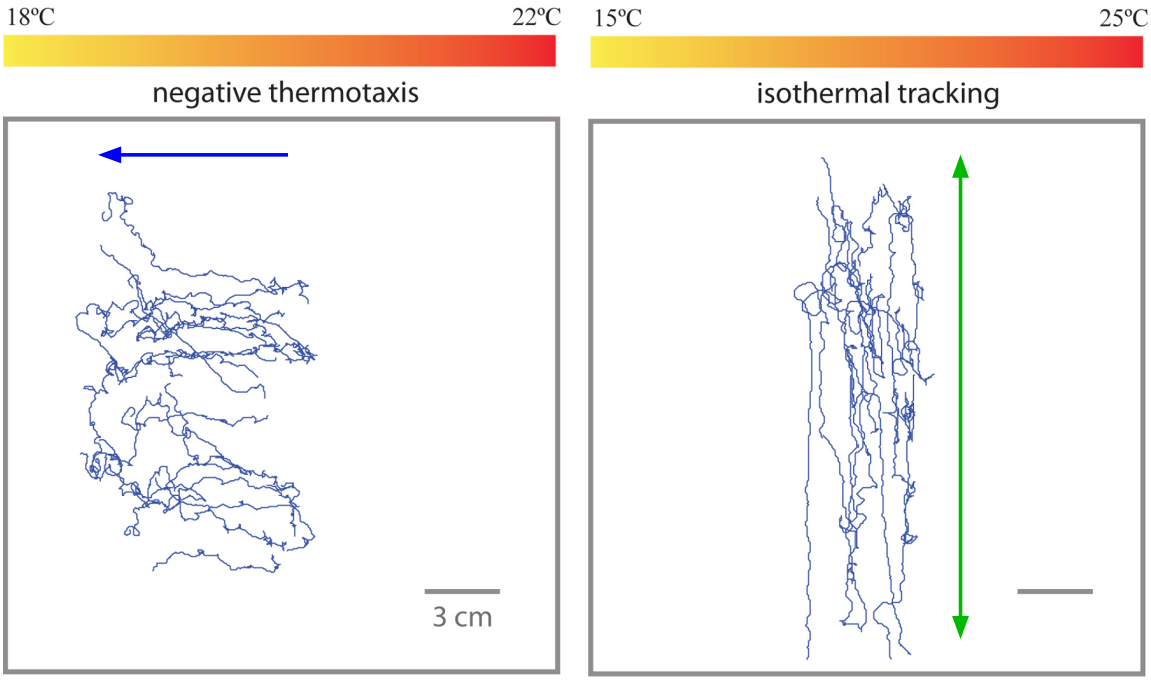


Figure 4.2: Population Thermotaxis (Ttx) Assays. Describes trajectories traced by 15 worms on agar plates in a temperature gradient. Left - Negative thermotaxis (cryophily) shown by worms having T_C = 15°C. Initial placement at 20°C, aggregate movement to a lower temperature results in a negative thermotaxis index (TI). Right - Worms at or near T_C, track isotherms, with a TI $\sim$ 0. Source [Goo14].

cell, it has bristle-like dendritic projections (refer Fig. 4.1) that are sensitive to temperature change, making it act like a contrast detector. Furthermore, it gates T_C information presynaptically to its downstream partner, the interneuron AIY, serving as memory (suppressed when $T \leq T_C$) [HCL$^+$18]. These functions are indispensable for thermotaxis, as killing the AFD renders the worm athermotactic. The AWC is a secondary thermosensor that shows a more stochastic response to temperatures further from T_C, while the response is suppressed at $T \sim T_C$ [BWT$^+$08]. Studies have implied the necessity of the AIY and AIZ first layer interneurons as their ablation leads to constitutive negative and positive thermotaxis respectively. They synapse onto the second layer command interneuron RIA, which connects to head motor neurons SMD and RMD for thermotaxis (Section 5.2). The circuit arrangement suggests that AIY and AIZ act in an antagonistic fashion, giving rise to the "two-drive model" (Fig. 4.3) [NSNM11].

While the aforementioned circuit is necessary, it is insufficient. It ignores a plethora of

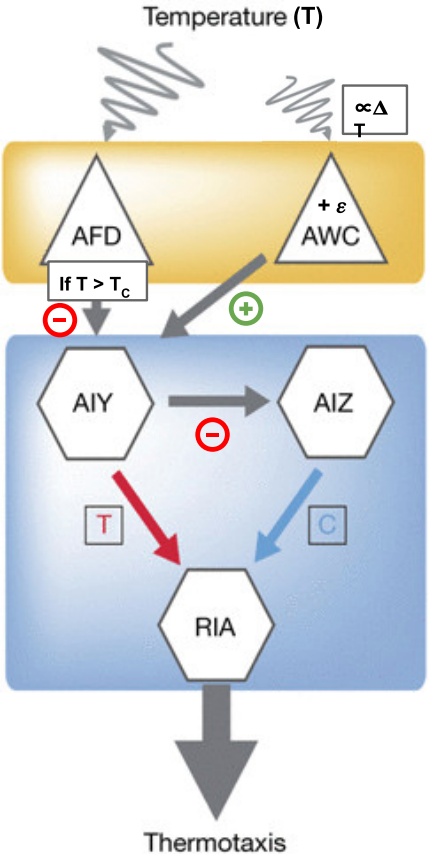


Figure 4.3: Updated "Two-Drive" Model for Thermotaxis in *C. elegans*.. Here, AIY and AIZ promote thermophily and cryophily respectively, via RIA. '+' and '-' labels indicate excitatory and inhibitory postsynaptic effects respectively. Therefore, when $T > T_C$, the former's activity decreases, and vice versa. At $T \sim T_C$, both drives are significantly reduced, and isothermal tracking predominates. Adapted from [NSNM11].

connections from interneurons that fine tune the worm's movement, and the motor neurons themselves, which are required for locomotion. Since a connectome is an organism's ultimate BNN, it is a sufficient circuit for its entire behavioral repertoire. By testing a model that incorporates as many necessary neurons and their neighbors (a neuron in a network that is isolated from its partners is equivalent to being deleted) as possible, one could hope to narrow the search to a minimal circuit and eventually interpret its mechanism. This is the aim of the author in the next section.

Chapter 5

An ANN Model of Thermotaxis

5.1 Artificial Neural Networks (ANN)

In Fig. 1.1, we saw the typical features of a neuron. For a neuron to transmit an electrical signal, it needs to experience a change in membrane potential, resulting in ion flow. This occurs due to postsynaptic potential differences contributed by sensory transduction or presynaptic neurons. This process can occur at every synapse. While synapses are usually found between axon terminals and dendrites, they can be formed between any region of the axon and an entire neuron, termed *en passant* synapses (predominant in *C. elegans*) [WSTB86]. The spatiotemporal mixing of these potentials is complex, but the change if sufficiently large, can result in a current that reaches a synapse. In a nutshell, artificial neurons are simplified versions of their biological counterparts (Figure 5.1).

Artificial neurons summarize all of the features of biological neurons succinctly:

$$a_k^l = \phi(\sum_{j=1}^{n}(a_j^{l-1}w_{jk}^l) + b_k^l) \tag{5.1}$$

where a_k^l is the activation of the k^{th} neuron in the l^{th} layer of an ANN. It is the result of a nonlinear transformation (ϕ) applied to the weighted (w) sum of activations of presynaptic neurons (indexed as j) plus an additive bias (b). The purpose of ϕ becomes relevant for chemical synapses. As discussed in Section 3.2, chemical synapses are sites of both excitatory (E, +ve) and inhibitory (I, -ve) postsynaptic potentials (PSP) which compete for determining the final PSP of the postsynaptic neuron. If not depolarized sufficiently, an action potential won't be generated and/or voltage gated ion channels at the next chemical

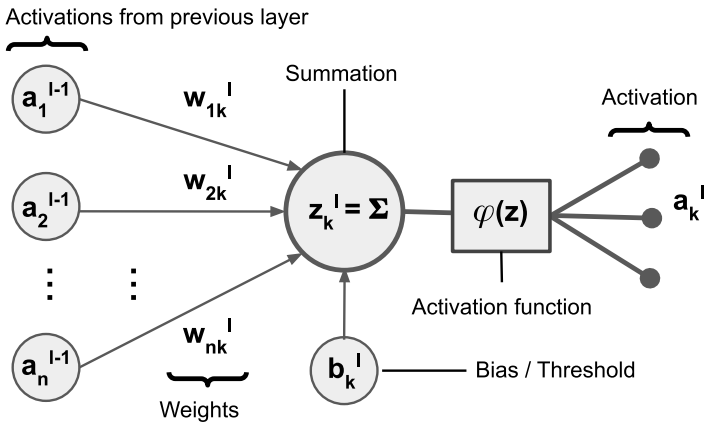


Figure 5.1: **Artificial Neuron**. Activations (**a**) are a result of current influx from presynaptic neurons. Weights (**w**) act as proxies for synapses. Weighted activations are summated (**z**), in a manner similar to the addition of postsynaptic potentials (PSP) in the cell body (see Fig. 1.1). The activation function (ϕ) transforms **z** non-linearly to yield the neuron's **a**, continuing the process. The bias or threshold (**b**) is an additional additive term to **z** that can alter the probability of ϕ giving a non-zero **a**.

synapse will fail to open, and so will neurotransmitter release. The bias term can be understood as the final PSP threshold for depolarization/neurotransmitter release. Weights can be positive/negative to indicate the nature of incoming PSPs, besides the number of synapses involved. When the summation of weighted activations of presynaptic neurons exceeds the bias term, $z > 0$ (Eqn. 5.1), the activation function should output a positive value. Else, it should output 0. The ReLU function is one such example (next section).

As indicated from Eqn. 5.1, an artificial neuron is rarely employed individually, rather, is a unit of a much larger *artificial neural network* (ANN, Figure 5.2):

ANNs come in many flavors. They are unified by their basic entity, the artificial neuron, and a *processing depth*, i.e., from input to output, there is a defined flow of activations such that every artificial neuron has all of its incoming weights associated with either the input or an activation. Such a hierarchy allows for layers of neurons to be formed, within which all neurons are activated simultaneously, in a *feed-forward* fashion as seen in Fig. 5.2. This normally prohibits neurons in deeper layers from sending activations to those before, within the same layer, or to themselves. Such connections are called *feedback-/recurrent connections*. They can exist, provided that recurrent activations are stored in

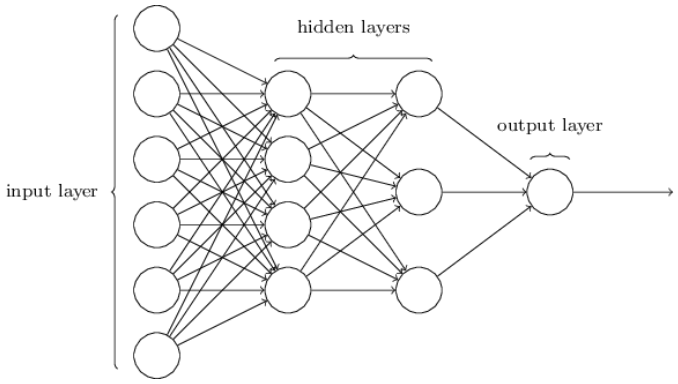


Figure 5.2: Simple Feed-forward ANN. This ANN has 4 layers, one input layer, one output layer with a single artificial neuron, and two 'hidden' layers with 4 and 3 artificial neurons respectively. Layers are demonstrated to be 'affine', i.e., fully connected, but can be made sparse by setting some weights to zero. Source [LeD].

memory and are utilized in a future *pass* (one complete input to output process). Additionally, feed-forward connections can skip layers to provide input deeper in the network, through *skip/residual connections*. Designing ANN architectures and methods to train their free parameters (weights and biases) to extract patterns from large datasets is the focus of the field of *Deep Learning* (DL). While a full discussion is beyond the scope of this work, it suffices to say that ANN variants such as CNNs, LSTMs, and ResNets have started to achieve human level performance on tasks such as image classification and forecasting in the last decade [Nie15].

5.2 Designing an ANN for Thermotaxis

Motivated by the successes of DL, one might ask if BNNs provide an encoding for behavior. If yes, could an ANN simulate the same? Recent experiments have shown that ANNs built like BNNs and trained to perform behaviors yield neurons with experimentally observed properties [TRT18][HSE18]. To design an ANN for the thermotaxis circuit, the author combined information from the two-drive model (Section 4.2) and combined it with previously studied navigation circuits in *C. elegans*. Chemotaxis and thermotaxis circuits share many neurons downstream as well as at the sensory level, hence the circuit by Gray et. al inspired the author to make a 4 hidden layer ANN [GHB05]. Further scrutiny of a candidate thermotaxis circuit came from Monte Carlo simulations of various recorded

movement patterns by Ikeda et. al., resulting in the identification of neurons responsible for reversal turns and curves, determined to be the most important components of thermotaxis behavior [ING+20]. Figure 5.3 summarizes the resulting ANN:

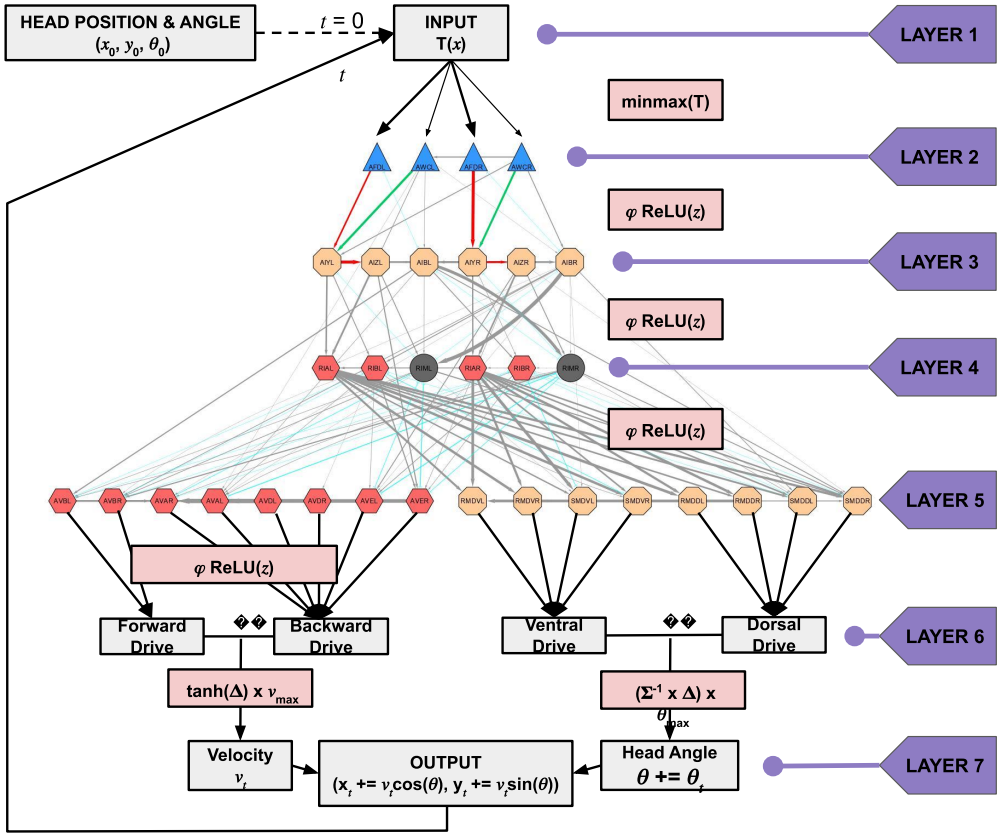


Figure 5.3: Proposed ANN for Thermotaxis. Subnetwork of the connectome (refer Fig. 3.2) for Ttx identified and organized in 4 layers to form a directed acyclic graph (DAG) corresponding to most likely signal flow hierarchy. Red and green arrows correspond to inhibitory and excitatory weights (refer Fig. 4.3 for individual detail). Grey boxes indicate engineered layers (3 in total) to derive meaningful information from the circuit. Pink boxes contain relevant transformations and activation functions between layers. The circuit is designed to operate at a frequency of one cycle per second, with input giving rise to output and vice-versa.

The ANN's construction keeps Ttx behavior in context. The 7 layer ANN was constructed as follows:

- **Layer 1 (Input)** - Based on the worm's initial position on a linear thermal gradient, the temperature (T) is calculated in °C, then compressed into a range of acceptable

values [0, 1] using a minmax function:

$$f(T) = \frac{T - T_{min}}{T_{max} - T_{min}}, Range = [0, 1] \tag{5.2}$$

where T_{min} and T_{max} are the minimum and maximum values that T can take respectively. This value is fed to all neurons in layer 2.

- **Layers 2 - 5 (Ttx circuit)** - The author constructed a *directed acyclic graph* (DAG) utilizing all previously mentioned experimental results. This graph has the property of having no feedback connections or cycles (discussed in the previous section), though unidirectional edges within layers are permitted. To do so, all electrical connections within layers were removed, followed by the removal of reciprocal feedback edges between layers. Within a layer, the weakest connection in a given loop was removed. Hence, information flows in the network only to the next layer, while having a depth hierarchy within each layer to ensure all neurons have defined input. A DAG is justified as the network is designed to have no memory apart from the temperature dependent weights between AFD and AIY, as well as T and AWC (further justification is provided when discussing the output layer below).
  Furthermore, rectified linear units (ReLU) activation functions were used for all layers:

$$\phi(z) = max(0, z) \tag{5.3}$$

  except layer 5, where two distinct groups of neurons are found: the red command interneurons for locomotion, and yellow polymodal neurons for head turns (Fig. 5.3). The latter form a terminal processing layer, and their outputs are taken as is. The biggest caveat is that the signs of weights, apart from those determined from the two-drive model, remain unknown. Hence, the author invoked the strategy of assigning all outgoing edges from interneurons as inhibitory (not shown in 5.3), which has been shown to yield computational results in best agreement with experimental data on locomotion [RK17].

- **Layer 6 (Drive)** - Based on the function and/or location of neurons in layer 5, layer 6 simply sums the outputs into four independent groups, termed 'drives'. They encode forward/backward movement and ventral/dorsal rotation respectively. Within each pair, the drives are antagonistic, hence a difference (Δ) of their sums is considered

to instruct the worm the extent to which it must execute a drive. To yield physically meaningful values, they are compressed with the functions:

$$v_t = tanh(\Delta) \cdot v_{max} \tag{5.4}$$

where v_t is the velocity of the worm, tanh takes any real valued input and squeezes it to (-1, 1), and v_{max} = 0.52 mms^{-1} is an experimental value for the maximum velocity of a worm for a second [HKD$^+$15].

$$\theta_t = \theta_{t-1} + (\Sigma^{-1} \cdot \Delta)\theta_{max} \tag{5.5}$$

Here, θ_t and θ_{t-1} are the current and previous head orientation angles w.r.t. the X axis, θ_{max} lies on [-π, π]. This is the maximum turn angle to the ventral or dorsal side, which is seen as anticlockwise or clockwise to the experimenter, if using the convention in Fig. 4.1a. The inverse of the sum (Σ^{-1}) term scales the drive on [-1, 1] if one of the drives is zero.

- **Layer 7 (Output)** - Finally, the drives from layer 6 are used to determine the next location visited by the worm:

$$(x_t, y_t) = (v_t cos(\theta_t), v_t sin(\theta_t)) \tag{5.6}$$

where (x_t, y_t) represents the new coordinates of the worm. The entire 7 layer network is defined for a single second of movement in the worm. Therefore, the process is iterated to obtain the worm's trajectory, represented by a feedback connection to layer 1.

5.3 Simulation and Results

To test the ANN's performance, the author designed a simple mock Ttx experiment protocol. All code was written in Python [VRDJ95] with the SciPy libraries NumPy [Oli06], pandas [M$^+$10], and NetworkX [HSS08] for array operations, data handling, and network preparation respectively. Plotting and visualizations were done with Matplotlib [Hun07] and Seaborn [WBO$^+$17] respectively:

1. **Gradient Preparation** - An artificial linear temperature gradient was established horizontally across a 20x20 cm^2 square, resulting in a 0.5°Ccm^{-1} increase ranging

from 15 - 25°C, (similar to Fig. 4.2R in Section 4.1) [Goo14]:

$$T(x) = \frac{x}{2(T_{max} - T_{min})) + T_{min}} \tag{5.7}$$

where x is the abscissa and T_{max} = 27°C and T_{min} = 12°C are considered to be the physiological limits of favorable temperature (Section 4.1).

2. **Hyperparameter Tuning** - Qualitative variables in the two-drive circuit (Fig. 4.3, Section 4.2) need to be quantified. AFD neurons were assigned an incoming weight of w = 10 while AWC neurons were assigned an incoming weight of w = 5 from morphological considerations (see Fig. 4.1a). Stochastic jitter ϵ was added to AWC taken from a Gaussian distribution with μ = 0 and σ = 0.1 (so that stochastic effects are small w.r.t. the input.)

3. **Initialization** - 3 groups of 15 'worms' each were encoded to have T_C representing cold (17°C), moderate (20°C), and hot (23°C) populations. Then, they were placed on the square at 3 positions corresponding to 23°C, 20°C, and 17°C respectively, along the Y axis midline. Since this setup is a computational one, interaction effects between the worms are nil and can hence be placed atop each other. Initial orientation angles are selected at random.

4. **Iteration** - The ANN is iterated repeatedly, for t = 3600s / 1hr. (see Fig. 5.3).

Figure 5.4 summarizes the results obtained.

The trajectories obtained cover a roughly spherical area. This is the result that would be expected from a 2D random walk, i.e., behavior in the absence of thermotaxis. Looking at the displacements of individual worms, there seems to be no preferred orientation for the green population, while the red and blue populations show some diagonal preference. This needs to be investigated in more detail. Reversing the gradient (to ensure no bias along the horizontal, not shown) yielded similar results.

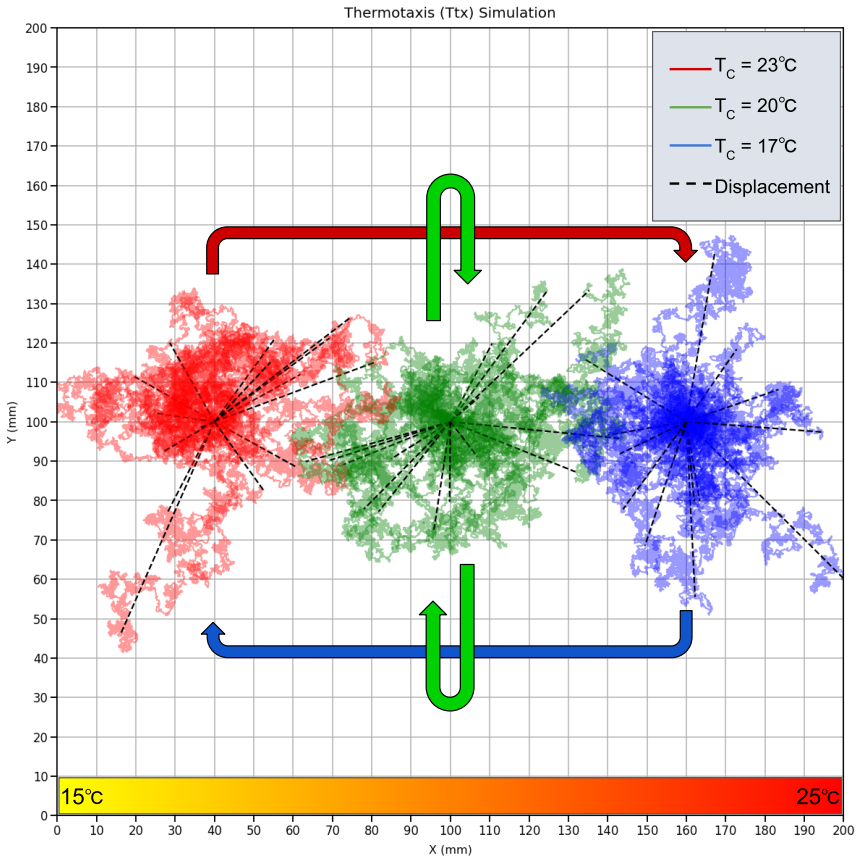


Figure 5.4: Results of Ttx Simulation. Red, green, and blue trajectories correspond to hot, moderate, and cold population worms. Dashed black lines denote displacement of individual worms over an hour. Red worms and blue worms are expected to exchange positions while green worms are expected to track isotherms along the Y axis. Instead, all three populations show random walk behavior.

Chapter 6

Discussion

The *C. elegans* connectome serves as a beacon to the field of connectomics. As an example of a primordial nervous system, it gives insights into how a 302 neuron nervous system could have evolved from a simpler nerve net, while increasing its connections by an order of magnitude. Despite reconstructions of the same being feasible, there is still uncertainty in some of its connections due to technological limitations and variability among worms. This is only expected to get worse as nervous systems become increasingly complex.

However, this complexity seems to follow certain rules to maximize efficiency. By analyzing the worm connectome as a network, it becomes clear that not all neurons are created equal. A handful of interneurons possess a disproportionately large share of connections, and are well connected to each other. These hubs play vital roles in coordinating the locomotion of the worm, as well as integrating information. Comparative studies with macroscale connectomes indicate similar structure, on a larger scale. This is a consequence of scale-freeness, which comes with a power law degree distribution. Moreover, the *C. elegans* connectome is a small world network, making it efficient for information flow. While small structural modules are expected from such a network, greedy algorithms fail to capture large anatomical compartments. This isn't unexpected, given the rudimentary organization of the worm connectome. Still, there is a definite sense of direction and signal flow in this connectome.

The connectome has served as a map for experimentalists to deduce behavior in the worm. Thermotaxis is a deceptively simple behavior which is not unique to *C. elegans*, but crucial

for its survival. Decades of experimentation has confirmed the necessary elements of the connectome responsible for this behavior, but the full mechanism remains unknown. One of the criticisms of connectomics has remained that the structure alone explains little about overlying function. However, artificial neural networks have shown us the importance of network structure in solving tasks previously believed to be inaccessible to computers.

Taken together, the author proposes that the future of connectomics and systems neuroscience requires us to start by solving a 'simple' problem of thermotaxis in the 'simple' nervous system of *C. elegans* by modelling an ANN on the connectome. While studies have shown that connectome inspired ANNs are capable of predicting individual neuron function, they all involve training the model to various extents. Since thermotaxis is an innate behavior, the author attempted to model thermotaxis with an untrained ANN, hypothesizing that the connectome is sufficient for the same.

The results obtained from the model are reminiscent of a 2D random walk, suggesting that the ANN constructed was unable to reproduce experimental observations in line with the two-drive model. However, this is just the first step, and the author acknowledges the following shortcomings. First and foremost, the model is constructed without knowing the appropriate signs of connection weights. While all interneuron outgoing edges were set to be inhibitory based on locomotory simulations, twenty other neurons are assumed to be excitatory in nature. This is unjustified. Other ANN experiments have taken into consideration neurotransmitter expression and receptor profiles for making educated guesses about the nature of each edge, and will be followed up by the author in the near future.

Secondly, the engineered layers 6 and 7 need to be trained to give desirable output. Doing so would not go against any of the principles of the author, as these layers are not naturally present in the connectome, but are abstractions. If the model shows thermotaxis behavior on training these layers, it would imply that the connectome has a representation of the input which can be mapped to the thermotaxis output. However, interpreting these layers and relating them to downstream neurons would be the next task. Other assumptions such as input weights, stochastic mixing in AWC, and the feed-forward nature of the model may be imperfect, but are nevertheless experimentally or theoretically motivated.

Once the model is able to show thermotaxis, it could be analyzed every second, to check neuronal activation for thermotaxis and cryotaxis respectively. If the representations in-

clude the well established two-drive model, it could serve as a guide to experimentalists to deduce the complete circuit. This interface between deep learning and neuroscience is exciting, and has promise to grow synergistically. In conclusion, there is much we still don't know about how the brain function depends on its architecture, and presented here is an attempt the author considers a worthwhile starting point.

Bibliography

[AHW+20] Z.F. Altun, L.A. Herndon, C.A. Wolkow, C. Crocker, R. Lints, and D.H. Hall, *Wormatlas*, 2002 - 2020.

[BKF18] Sophie H. Bennett, Alastair J. Kirby, and Gerald T. Finnerty, *Rewiring the connectome: Evidence and effects*, Neuroscience and Biobehavioral Reviews **88** (2018), 51–62.

[BWT+08] David Biron, Sara Wasserman, James H. Thomas, Aravinthan D. T. Samuel, and Piali Sengupta, *An olfactory neuron responds stochastically to temperature and modulates Caenorhabditis elegans thermotactic behavior*, Proceedings of the National Academy of Sciences **105** (2008), no. 31, 11002–11007 (en), Publisher: National Academy of Sciences Section: Biological Sciences.

[CEbAH00] Reuven Cohen, Keren Erez, Daniel ben Avraham, and Shlomo Havlin, *Resilience of the Internet to Random Breakdowns*, Physical Review Letters **85** (2000), no. 21, 4626–4628, Publisher: American Physical Society.

[CH03] Reuven Cohen and Shlomo Havlin, *Scale-Free Networks Are Ultrasmall*, Physical Review Letters **90** (2003), no. 5, 058701, Publisher: American Physical Society.

[CNM04] Aaron Clauset, M. E. J. Newman, and Cristopher Moore, *Finding community structure in very large networks*, Physical Review E **70** (2004), no. 6, 066111, arXiv: cond-mat/0408187.

[Dhp19] Dhp1080, *File:Neuron.svg - Wikimedia Commons*, 2019.

[Emm15] Scott W. Emmons, *The beginning of connectomics: a commentary on White et al. (1986) 'The structure of the nervous system of the nematode Caenorhab-*

ditis elegans', Philosophical Transactions of the Royal Society B: Biological Sciences **370** (2015), no. 1666.

[GHB05] J. M. Gray, J. J. Hill, and C. I. Bargmann, *A circuit for navigation in Caenorhabditis elegans*, Proceedings of the National Academy of Sciences **102** (2005), no. 9, 3184–3191 (en).

[Gli06] Mitch Glickstein, *Golgi and Cajal: The neuron doctrine and the 100th anniversary of the 1906 Nobel Prize*, Current Biology **16** (2006), no. 5, R147–R151 (English), Publisher: Elsevier.

[Goo14] Miriam B. Goodman, *Thermotaxis navigation behavior*, pp. 1–10, 2014.

[HCL+18] Josh D. Hawk, Ana C. Calvo, Ping Liu, Agustin Almoril-Porras, Ahmad Aljobeh, María Luisa Torruella-Suárez, Ivy Ren, Nathan Cook, Joel Greenwood, Linjiao Luo, Zhao-Wen Wang, Aravinthan D. T. Samuel, and Daniel A. Colón-Ramos, *Integration of Plasticity Mechanisms within a Single Sensory Neuron of C. elegans Actuates a Memory*, Neuron **97** (2018), no. 2, 356–367.e4 (en).

[HG08] Mark D. Humphries and Kevin Gurney, *Network 'Small-World-Ness': A Quantitative Method for Determining Canonical Network Equivalence*, PLoS ONE **3** (2008), no. 4.

[HKD+15] Jeong-Hoon Hahm, Sunhee Kim, Race DiLoreto, Cheng Shi, Seung-Jae V. Lee, Coleen T. Murphy, and Hong Gil Nam, *C. elegans maximum velocity correlates with healthspan and is maintained in worms with an insulin receptor mutation*, Nature Communications **6** (2015), no. 1, 8919 (en), Number: 1 Publisher: Nature Publishing Group.

[Hob10] O. Hobert, *Neurogenesis in the nematode Caenorhabditis elegans*, 2010.

[HSE18] Martin Haesemeyer, Alexander F Schier, and Florian Engert, *Convergent temperature representations in artificial and biological neural networks*, bioRxiv (2018), 390435.

[HSS08] Aric A. Hagberg, Daniel A. Schult, and Pieter J. Swart, *Exploring network structure, dynamics, and function using networkx*, Proceedings of the 7th

Python in Science Conference (Pasadena, CA USA) (Gaël Varoquaux, Travis Vaught, and Jarrod Millman, eds.), 2008, pp. 11 – 15.

[Hun07] John D Hunter, *Matplotlib: A 2d graphics environment*, Computing in science & engineering **9** (2007), no. 3, 90–95.

[ING+20] Muneki Ikeda, Shunji Nakano, Andrew C. Giles, Linghuan Xu, Wagner Steuer Costa, Alexander Gottschalk, and Ikue Mori, *Context-dependent operation of neural circuits underlies a navigation behavior in Caenorhabditis elegans*, Proceedings of the National Academy of Sciences **117** (2020), no. 11, 6178–6188 (en), Publisher: National Academy of Sciences Section: Biological Sciences.

[KBSK20] Konrad P. Kording, Gunnar Blohm, Paul Schrater, and Kendrick Kay, *Appreciating the variety of goals in computational neuroscience*, arXiv:2002.03211 [q-bio] (2020), arXiv: 2002.03211.

[Kel15] Susan Keller, *The lateral giant interneuron input and output - Motor Neurons*, April 2015, Library Catalog: www.alpfmedical.info.

[KL07] Narayanan Kasthuri and Jeff W. Lichtman, *The rise of the 'projectome'*, Nature Methods **4** (2007), no. 4, 307–308 (en), Number: 4 Publisher: Nature Publishing Group.

[KW78] Irving Kupfermann and Klaudiusz R. Weiss, *The command neuron concept*, Behavioral and Brain Sciences **1** (1978), no. 1, 3–10 (en), Publisher: Cambridge University Press.

[Lea] Mark Leaver, *C. elegans*.

[LeD] Erin LeDell, *ledell/sldm4-h2o*, Library Catalog: github.com.

[M+10] Wes McKinney et al., *Data structures for statistical computing in python*, Proceedings of the 9th Python in Science Conference, vol. 445, Austin, TX, 2010, pp. 51–56.

[MdFCC07] Julian J. McAuley, Luciano da Fontoura Costa, and Tibério S. Caetano, *Rich-club phenomenon across complex network hierarchies*, Applied Physics Letters **91** (2007), no. 8, 084103, Publisher: American Institute of Physics.

[New10] M. E. J. Newman, *Networks: an introduction*, Oxford University Press, Oxford; New York, 2010.

[Nie15] Michael A. Nielsen, *Neural Networks and Deep Learning*, Determination Press, 2015 (en).

[noa] *Computational neuroscience - Latest research and news | Nature*.

[noa20] *List of animals by number of neurons*, June 2020, Page Version ID: 961059727.

[NSNM11] Yukuo Nishida, Takuma Sugi, Mayu Nonomura, and Ikue Mori, *Identification of the AFD neuron as the site of action of the CREB protein in Caenorhabditis elegans thermotaxis*, EMBO reports **12** (2011), 855–62.

[OHN+14] Seung Wook Oh, Julie A. Harris, Lydia Ng, Brent Winslow, Nicholas Cain, Stefan Mihalas, Quanxin Wang, Chris Lau, Leonard Kuan, Alex M. Henry, Marty T. Mortrud, Benjamin Ouellette, Thuc Nghi Nguyen, Staci A. Sorensen, Clifford R. Slaughterbeck, Wayne Wakeman, Yang Li, David Feng, Anh Ho, Eric Nicholas, Karla E. Hirokawa, Phillip Bohn, Kevin M. Joines, Hanchuan Peng, Michael J. Hawrylycz, John W. Phillips, John G. Hohmann, Paul Wohnoutka, Charles R. Gerfen, Christof Koch, Amy Bernard, Chinh Dang, Allan R. Jones, and Hongkui Zeng, *A mesoscale connectome of the mouse brain*, Nature **508** (2014), no. 7495, 207–214.

[Oli06] Travis E Oliphant, *A guide to numpy*, vol. 1, Trelgol Publishing USA, 2006.

[Ram19] Venkatakrishnan Ramaswamy, *An Algorithmic Barrier to Neural Circuit Understanding*, bioRxiv (2019), 639724 (en), Publisher: Cold Spring Harbor Laboratory Section: New Results.

[RK17] Franciszek Rakowski and Jan Karbowski, *Optimal synaptic signaling connectome for locomotory behavior in Caenorhabditis elegans: Design minimizing energy cost*, PLOS Computational Biology **13** (2017), no. 11, e1005834 (en), Publisher: Public Library of Science.

[SMO+03] Paul Shannon, Andrew Markiel, Owen Ozier, Nitin S. Baliga, Jonathan T. Wang, Daniel Ramage, Nada Amin, Benno Schwikowski, and Trey Ideker,

Cytoscape: a software environment for integrated models of biomolecular interaction networks, Genome Research **13** (2003), no. 11, 2498–2504 (eng).

[Spo10] Olaf Sporns, *Connectome*, Scholarpedia **5** (2010), no. 2, 5584 (en).

[TRT18] Fabian David Tschopp, Michael B. Reiser, and Srinivas C. Turaga, *A Connectome Based Hexagonal Lattice Convolutional Network Model of the Drosophila Visual System*, arXiv:1806.04793 [cs, q-bio] (2018), arXiv: 1806.04793.

[TVA+13] Emma K. Towlson, Petra E. Vértes, Sebastian E. Ahnert, William R. Schafer, and Edward T. Bullmore, *The Rich Club of the C. elegans Neuronal Connectome*, Journal of Neuroscience **33** (2013), no. 15, 6380–6387 (en), Publisher: Society for Neuroscience Section: Articles.

[VCP+11] Lav R. Varshney, Beth L. Chen, Eric Paniagua, David H. Hall, and Dmitri B. Chklovskii, *Structural Properties of the Caenorhabditis elegans Neuronal Network*, PLOS Computational Biology **7** (2011), no. 2, e1001066 (en), Publisher: Public Library of Science.

[vdHBS16] Martijn P. van den Heuvel, Edward T. Bullmore, and Olaf Sporns, *Comparative Connectomics*, Trends in Cognitive Sciences **20** (2016), no. 5, 345–361 (en).

[VRDJ95] Guido Van Rossum and Fred L Drake Jr, *Python reference manual*, Centrum voor Wiskunde en Informatica Amsterdam, 1995.

[WBO+17] Michael Waskom, Olga Botvinnik, Drew O'Kane, Paul Hobson, Saulius Lukauskas, David C Gemperline, Tom Augspurger, Yaroslav Halchenko, John B. Cole, Jordi Warmenhoven, Julian de Ruiter, Cameron Pye, Stephan Hoyer, Jake Vanderplas, Santi Villalba, Gero Kunter, Eric Quintero, Pete Bachant, Marcel Martin, Kyle Meyer, Alistair Miles, Yoav Ram, Tal Yarkoni, Mike Lee Williams, Constantine Evans, Clark Fitzgerald, Brian, Chris Fonnesbeck, Antony Lee, and Adel Qalieh, *mwaskom/seaborn: v0.8.1 (september 2017)*, September 2017.

[WS98] Duncan J. Watts and Steven H. Strogatz, *Collective dynamics of 'small-world'*

networks, Nature **393** (1998), no. 6684, 440–442 (en), Number: 6684 Publisher: Nature Publishing Group.

[WSTB86] John Graham White, Eileen Southgate, J. N. Thomson, and Sydney Brenner, *The structure of the nervous system of the nematode Caenorhabditis elegans*, Philosophical Transactions of the Royal Society of London. B, Biological Sciences **314** (1986), no. 1165, 1–340, Publisher: Royal Society.

[XJL+20] C. Shan Xu, Michal Januszewski, Zhiyuan Lu, Shin-ya Takemura, Kenneth J. Hayworth, Gary Huang, Kazunori Shinomiya, Jeremy Maitin-Shepard, David Ackerman, Stuart Berg, Tim Blakely, John Bogovic, Jody Clements, Tom Dolafi, Philip Hubbard, Dagmar Kainmueller, William Katz, Takashi Kawase, Khaled A. Khairy, Laramie Leavitt, Peter H. Li, Larry Lindsey, Nicole Neubarth, Donald J. Olbris, Hideo Otsuna, Eric T. Troutman, Lowell Umayam, Ting Zhao, Masayoshi Ito, Jens Goldammer, Tanya Wolff, Robert Svirskas, Philipp Schlegel, Erika R. Neace, Christopher J. Knecht, Chelsea X. Alvarado, Dennis A. Bailey, Samantha Ballinger, Jolanta A. Borycz, Brandon S. Canino, Natasha Cheatham, Michael Cook, Marisa Dreher, Octave Duclos, Bryon Eubanks, Kelli Fairbanks, Samantha Finley, Nora Forknall, Audrey Francis, Gary Patrick Hopkins, Emily M. Joyce, SungJin Kim, Nicole A. Kirk, Julie Kovalyak, Shirley A. Lauchie, Alanna Lohff, Charli Maldonado, Emily A. Manley, Sari McLin, Caroline Mooney, Miatta Ndama, Omotara Ogundeyi, Nneoma Okeoma, Christopher Ordish, Nicholas Padilla, Christopher Patrick, Tyler Paterson, Elliott E. Phillips, Emily M. Phillips, Neha Rampally, Caitlin Ribeiro, Madelaine K. Robertson, Jon Thomson Rymer, Sean M. Ryan, Megan Sammons, Anne K. Scott, Ashley L. Scott, Aya Shinomiya, Claire Smith, Kelsey Smith, Natalie L. Smith, Margaret A. Sobeski, Alia Suleiman, Jackie Swift, Satoko Takemura, Iris Talebi, Dorota Tarnogorska, Emily Tenshaw, Temour Tokhi, John J. Walsh, Tansy Yang, Jane Anne Horne, Feng Li, Ruchi Parekh, Patricia K. Rivlin, Vivek Jayaraman, Kei Ito, Stephan Saalfeld, Reed George, Ian Meinertzhagen, Gerald M. Rubin, Harald F. Hess, Louis K. Scheffer, Viren Jain, and Stephen M. Plaza, *A Connectome of the Adult Drosophila Central Brain*, bioRxiv (2020),

2020.01.21.911859 (en), Publisher: Cold Spring Harbor Laboratory Section: New Results.

[Xu19] Nan Xu, *Deep phenotyping in C. elegans*, 16 (en).