Computational Neuroscience Course Description & Prerequisites

In 2005 while serving as deputy provost at Brown I developed and then implemented a computational model of David Mumford and Tai Sing Lee’s “Hierarchical Bayesian Inference in the Visual Cortex” using Bayesian Networks running on the department’s grid-computing cluster. It was a fascinating introduction to what was for me a new field of scientific inquiry that has guided my professional career ever since.

I started teaching “Computational Models of the Neocortex” at Stanford while on sabbatical in 2006 and continued as a lecturer and then consulting faculty at Stanford while at Google. The course evolved considerably over the years while the title remained the same and the focus shifted to how our understanding of the brain is yielding insights into machine learning and AI and how industrial-scale high-performance computing might expedite such discoveries.

Neurons in the primate neocortex support distributed computations involving both dense local connections and high-speed reciprocal white matter tracts that span the length and breadth of the brain. Recurrent neural networks, both biological and artificial, are powerful in part because they can store the intermediate results from earlier (recurrent) passes through the network.

Over the past decade, advanced memory technologies have substantially improved the scope and performance of artificial neural network technologies. The following four technologies have been particularly influential and were derived in part from our understanding of biological memory systems: Long-Short Term Memory LSTM, Neural Turing Machines NTM, Differential Neural Computer DNC, and Dense Continuous (or so-called “modern”) Hopfield Networks.

The inspiration for these memory technologies came from progress in understanding how basal ganglia, thalamus, hippocampus, and the prefrontal and entorhinal cortex work together in solving cognitive tasks. Locally sensitive hashing (LSH) and simultaneous localization and mapping SLAM are just two of many problems for which advances in neural recording technologies have yielded more efficient solutions from studying, respectively, the olfactory bulb in fruit flies and the network of place and grid cells in mice.

Coming up with an algorithm from scratch for any nontrivial computing problem would be difficult, if not impossible, for modern programmers without the effort of generations of hackers and software engineers. Think back to the development of the C programming language, UNIX, and other tools including grep, sed, awk, etc. developed at Bell Labs in the ’70s. Programmers today live in a rich software ecosystem full of accessible software. The analogy to a biological ecosystem is apt in this context.

Only recently have methods for tracing neural circuits and recording from neurons at scale become available. My team at Google helped to accelerate this trend and now universities and companies including DeepMind are using such data to discover new algorithms and develop new technologies. The proposed course is intended to introduce students to the tools and methodologies that are driving this approach to building advanced AI systems. The following four slides illustrate our pedagogical approach.
Sample Lecture Illustrating the Course Pedagogical Approach

Figure 1: The graphic shows a neural network known as a Transformer familiar to researchers working in machine learning. The recommended prerequisites for the class include familiarity with the basic subcircuits from which Transformers are constructed which will be reviewed in an earlier lecture. This slide describes the function of these subcircuits (see below) and how they contribute to the overall function of the network when applied to an NLP task such as machine translation or question answering. Initially, students are told to ignore the components S and M comprising the circular flow of information between an organism and its environment referred to as the perception-action cycle. The remainder of the lecture discusses how these same basic network components can be applied in agent architectures using algorithmic insights from neuroscience. This short summary will be considerably expanded in the lecture.

1. Encoding: For our purpose, an encoder maps one representation into a second more compact representation in a latent space. A latent space is a representation of compressed data in which similar data points are closer together in space.
2. Positional Encoding: Words derive meaning from their relative positions much like the relative positions of atoms in a protein molecule determine its functional properties.
3. Self Attention: You can think of a sentence as a graph in which words are nodes and edges correspond to semantic relationships between the words in the sentence. Then self-attention amounts to learning a relation-labeled relevance-weighted-edge graph.
4. Normalization and Skip Connections – also known as residual neural networks – are used to avoid the problem of vanishing gradients and mitigate the saturation problem in which adding more layers to a deep neural network leads to higher training error.
5. Masking and Positional Readjustment: During supervised learning, the target sequence is made visible to the decoder one word at a time to mimic natural sequence production.
6. Encoder-decoder attention: The decoder encoder-decoder attention layer combines the output of the encoder stack with the preceding self-attention layer in the decoder stack.
Figure 2: This slide revisits the transformer architecture from the previous slide mapping its different component artificial neural networks to biological counterparts that perform similar functions. Shown in the three insets: (A) the thalamus providing reciprocal connections between cortical areas and subcortical circuits including the hippocampus, and between the cortex and sensory systems including vision and audition; (B) the hippocampal complex (HPC) and its connections to the thalamus and entorhinal cortex; and (C) the basal ganglia (BG) with connections to the prefrontal cortex (PFC) and striatum by way of the thalamus. The next slide highlights the interplay between BG, PFC, and HPC illustrating how insights from cognitive neuroscience have led to significant improvements in reinforcement learning introduced in Matt Botvinick’s video below. The diagrams in the three insets will be described briefly in this lecture with a focus on their function. They figure prominently in the readings and subsequent discussions in class.

O’Reilly on Prefrontal Cortex, Basal Ganglia & Working Memory (URL)
Matt Botvinick on Reinforcement Learning & Meta-Learning (URL)
Figure 3: Until recently, the direct and indirect pathways were considered the primary pathways supporting recurrent action selection in the basal ganglia. We now know that the subthalamic nucleus (STN) is a major component of the basal ganglia which acts as the third hyper direct pathway, so named because it receives input directly from the frontal cortex and sends excitatory projections directly to the basal ganglia bypassing the striatum altogether. This area has been shown to become more active with increasing demands for response inhibition and appears to be implicated in how agents learn to learn. For more detail, see Matt Botvinick’s invited talk in the Stanford instantiation of my class and his slides for more on neuroscience-inspired AI at Deepmind. Randy OReilly’s Computational Cognitive Neuroscience and Bear, Connors, and Paradiso’s Neuroscience: Exploring the Brain are class secondary references.

Ressler on Amygdala, Olfaction, Hippocampus & Sensory Cortex (URL)
Greg Wayne on the Imitation of Intelligent Interactive Agents (URL)
Figure 4: This is a refinement of Figure 2 in which I've replaced the insets describing the structure and function of different neural circuits with artificial neural network technologies implementing key functions derived from our understanding of the neural circuits. In class, I would take the time and provide additional slides to explain in detail how we address the problem of partial observability to compress and embed state vectors using self-attention and graph learning in order to support action selection, and how we rely on positional encoding to leverage what we know about how the striatum, thalamus, entorhinal and prefrontal cortex are organized in order to facilitate coordinated distributed computation. I will also describe how we plan to employ a neural network model of place and grid cells to compactly encode reachable (sensorimotor) manifolds in high-dimensional dynamical-system state spaces amenable to interpolation, and dense associative Hopfield networks and attractor dynamics for exploiting episodic memory in zero or few-shot learning. The basic themes, applications, models, and proposed solutions mentioned above will be revisited throughout the semester to focus on class discussions and projects.

Artificial Agents Navigating with Grid-like Representations (URL)
Sherman Rethinking Transthalamic Corticocortical Circuitry (URL)
Prerequisites for New Computational Neuroscience Course

The prerequisites for the class I taught at Stanford and propose to teach at Brown are described in some detail in this document. The material is divided into two main sections with the first describing the basic machine learning and artificial neural network concepts assumed in this class, and the second providing a glimpse of how the class leverages insights from neuroscience in developing advanced machine learning technologies.

The latter should be accessible to students who are comfortable with the material in the first section. Those students who also have a background in neuroscience are particularly well prepared to benefit from the class, and in particular to join teams comprised of both CS and NS students working on final projects. At Stanford, the course attracted both graduate and undergraduate students, many of whom were interested in learning about the sort of applied research going on at Google Brain and DeepMind, both of which I have been involved with and helped students get internships and find jobs at.

For comparison to prerequisites for the class at Stanford, CS224 basically walks through the history of NLP and related machine learning methods including RNNs and Transformers. It doesn't cover CNNs, though don't ask me why since it is taught by Chris Manning who uses CNNs extensively in his NLP research. In contrast, CS231 emphasizes computer vision and spends substantial time covering convolutional networks. I've also had a number of Stanford students who have taken CS229 which provides the mathematical background – statistics, matrix algebra, and differential calculus, for understanding the neural network models discussed in class.

As for Brown courses that appear to provide the background for the proposed course, Ritambhara Singh’s CSCI1470 offered in the spring semester seems to be the best fit for its extensive coverage of Deep learning-based methods including convolutional neural networks, recurrent neural networks, autoencoders, and their applications. As a possible alternative, Srinath Sridhar and James Tompkin teach CSCI1430 Computer Vision in the fall and spring semesters respectively. Their course covers CNNs as well as foundational image processing including sampling, aliasing, and Fourier methods, along with related mathematical and computer science background.

Ellie Pavlick’s CSCI2952-1 Language Processing in Humans and Machines and Eugene Charniak’s CSCI242 Statistical Models in Natural-Language Processing coupled with a background in machine learning should suffice as prerequisites. Eugene’s Introduction to Deep Learning works as an aid for students who need help getting up to speed for class projects. In addition, Both Amy Greenwald and Michael Littman have taught classes that might provide the necessary background.

There are no required textbooks for the course. All the study materials will be available on the course website. If a student wishes to dig deeper into basic neuroscience and neuroanatomy, we recommend the textbook by Mark Bear, Barry Connors, and Michael Paradiso [1]. We will also reference excerpts from the textbook on computational cognitive neuroscience by Randall O’Reilly, Yuko Munakata, Michael Frank, and Thomas Hazy. However, this textbook is available online along with additional resources.


Sample Projects and Papers Involving CS379C Students

The [Stanford] Programmers Apprentice Project (URL)

The original programmer's apprentice project started at MIT in 1987. A version of the project borrowing the same name was started at Stanford in 2018 as an aspirational goal for students taking CS379C Computational Models of the Neocortex. The Stanford project has focused on building interactive digital assistants by leveraging recent advances in machine learning, artificial intelligence, and cognitive and systems neuroscience. The course features invited lectures by experts in these fields and attracts students from a range of disciplines. In several of the years following the end of class, selected students were invited to contribute to a technical report summarizing the material covered and lessons learned during the quarter. For each of the last three years, the class focused on two or three recent papers emphasizing important dimensions of the programmer's apprentice architecture. […]

Neurobot Collaboration with Henry Markram's Lab (URL)

Neurobot is a collection of tools for analyzing connectomic datasets and the HHMI Janelia Farm seven-column-medulla dataset. The original inspiration came from experimenting with Pawel Dlotko's Neurotop software which implements a class of simplicial complex called a directed flag complex described in a paper published on arXiv. The Dlotko et al paper is self-contained in terms of explaining the relevant mathematics. Still, you might want to look at David Cox's excellent primer on clique topology for a painless introduction.

In addition to specialized graph and visualization tools, Neurobot includes a convolution operator that applies Dlotko's code to compute topologically invariant properties of the subgraphs embedded in spherical subvolumes as defined by diameter and stride parameters. These properties are used to construct local feature vectors and classify regions of the connectome graph. This notebook introduces the reader to some of the most useful tools by demonstrating a typical workflow for analyzing the Janelia dataset.


Technology Prospects and Investment Opportunities for Scalable Neuroscience (URL)

“Two major initiatives to accelerate research in the brain sciences have focused attention on developing a new generation of scientific instruments for neuroscience. These instruments will be used to record static (structural) and dynamic (behavioral) information at unprecedented spatial and temporal resolution and report that information in a form suitable for computational analysis. We distinguish between recording — taking measurements of individual cells and the extracellular matrix — and reporting — transcoding, packaging, and transmitting the resulting information for subsequent analysis — as these represent very different challenges as we scale the relevant technologies to support simultaneously tracking the many neurons that comprise neural circuits of interest. We investigate a diverse set of technologies with the purpose of anticipating their development over the span of the next 10 years and categorizing their impact in terms of short [1-2 years], medium [2-5 years], and longer-term [5-10 years] deliverables.”

Inferring Mesoscale Models of Neural Computation (URL)

“Recent years have seen dramatic progress in the development of techniques for measuring the activity and connectivity of large populations of neurons in the brain. However, as these techniques grow ever more powerful—allowing us even to contemplate measuring every neuron in the entire brain—a new problem arises: how do we make sense of the mountains of data that these techniques produce? Here, we argue that the time is ripe for building an intermediate or "mesoscale" computational theory that can bridge between single-cell (microscale) accounts of neural function and behavioral (macroscale) accounts of animal cognition and environmental complexity. Just as digital accounts of computation in conventional computers abstract away the non-essential dynamics of the analog circuits that implement gates and registers, so too computational accounts of animal cognition can afford to abstract from the non-essential dynamics of neurons. We argue that the geometry of neural circuits is essential in explaining the computational limitations and technological innovations inherent in biological information processing. We propose a blueprint for how to employ tools from modern machine learning to automatically infer a satisfying mesoscale account of neural computation that combines functional and structural data, with an emphasis on learning and exploiting regularity and repeating motifs in neuronal circuits. Rather than suggest a specific theory, we present a new class of scientific instruments that can enable neuroscientists to design, propose, implement and test mesoscale theories of neural computation.”

Biological Blueprints for Next Generation AI Systems (URL)

“Diverse subfields of neuroscience have enriched artificial intelligence for many decades. With recent advances in machine learning and artificial neural networks, many neuroscientists are partnering with AI researchers and machine learning experts to analyze data and construct models. This paper attempts to demonstrate the value of such collaborations by providing examples of how insights derived from neuroscience research are helping to develop new machine learning algorithms and artificial neural network architectures. We survey the relevant neuroscience necessary to appreciate these insights and then describe how we can translate our current understanding of the relevant neurobiology into algorithmic techniques and architectural designs. Finally, we characterize some of the major challenges facing current AI technology and suggest avenues for overcoming these challenges that draw upon research in developmental and comparative cognitive neuroscience.”
Examples of Audio Visual Materials Used in Class

Structural Connectomics & Interactive Brain Atlases

Allen Atlas Mouse Brain Map Coronal View [URL](#)
Drosophila Melanogaster Atlas Coronal View [URL](#)
Connectome of an Adult Drosophila Brain [URL VIDEO](#)
Finch Basal Ganglia Flood-filling Networks [URL VIDEO](#)

Theoretical Computer Scientists & Brain Science

Christos Papadimitriou and Santosh Vempala
Computational and Learning with Assemblies of Neurons

Leslie Valiant
Valiant’s Model of Memory and Association in the Brain

Sanjoy Dasgupta
Neural Algorithm for Fundamental Computing Problem