Summary of Cognitive Computational Neuroscience Conference 2022

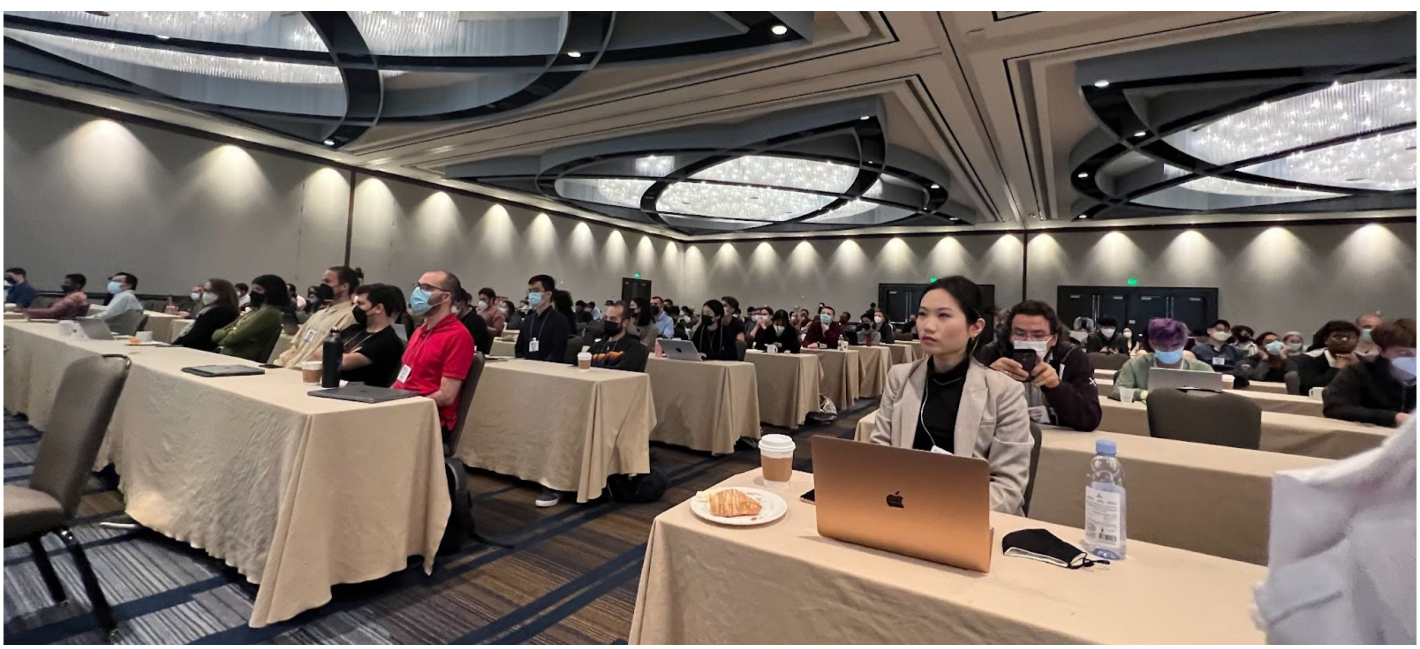
Emily Meschke | Helen Wills Neuroscience Institute, UC Berkeley | @MeschkeEmily

Sevan Harootonian | Department of Psychology, Princeton University | @harootonian

The Cognitive Computational Neuroscience (CCN) conference held August 25-28 in San Francisco brought together researchers from related disciplines to discuss how to build and validate computational models of human behavior. CCN took an explicitly interdisciplinary approach to address this high-level goal, inviting researchers from cognitive science, artificial intelligence, and neuroscience to contribute to the conference. However, for researchers to effectively communicate across these fields, CCN had to link results across multiple levels of abstraction, from neural activity to complex behaviors. The conference organizers took multiple measures to encourage collaborations and discussions across these three fields.

One way that CCN connected work from cognitive science, artificial intelligence, and neuroscience was by unifying research under a common framework. In this framework, intelligent behavior is modeled as an agent interacting with an environment by making observations, constructing beliefs, and performing actions. While all three disciplines are interested in some aspects of this framework, they tend to operate at different levels of abstraction. The field of cognitive science tends to focus on understanding how the agent constructs beliefs to take statistically well-informed actions. The field of artificial intelligence tends to focus on building models that achieve human-like levels of performance across perception, cognition, and behavior. The field of neuroscience tends to focus on collecting data from the brain to inform theories of how and where information is represented in the brain. Nonetheless, at CCN, all disciplines were brought together under the same goal of modeling human perception, cognition, and behavior.

Another way that CCN encouraged interdisciplinary collaboration was through the types of events that were hosted. The conference started with Mind Matching, presenting an opportunity for researchers to meet and discuss their shared research interests. The keynote talks were thematically organized, centering discussions around how the different fields approach the same question or problem and lastly, the generative adversarial collaborations (GACs) brought together numerous perspectives from experts while also encouraging contributions from audience members.



Perception
Before an agent can decide how to interact with the world, the agent must be able to *perceive the state of the world*. This aspect of the framework was tackled by the first GAC, which asked how data collected from the

brain should be used to inform models of visual perception. A key theme discussed in this GAC was the biological plausibility of deep artificial neural networks (ANNs). Most speakers agreed that ANNs should display some key characteristics of the brain system whose functions they are implementing. However, the level at which ANNs should mirror biological systems was somewhat disputed.

Dr. Joel Zylberberg from York University took a strong stance in favor of biologically plausible architecture. He argued that building ANNs with the same circuit structure and connectivity is the only way to achieve human-like intelligence with ANNs. This argument also extended to the input and output of the models. While most models of the human visual system receive a 2D matrix of pixel values as the input image, the human visual system receives a signal from photoreceptor cells in the retina. Other work has constructed biophysical models of the transduction pathway, mathematically describing how photoreceptors convert particles of light into a neural signal. Dr. Zylberberg suggested first performing these computations on the input images and then feeding this signal to the models of the human visual system. He argued that this process would help to resolve one of the largest problems with ANNs – their lack of robustness to different image sets with different visual tendencies.

Dr. Kalanit Grill-Spector from Stanford University argued for a higher-level, more abstracted version of biological plausibility. Her lab uses functional magnetic resonance imaging (fMRI) to map functional representations across the human visual pathway. She noted that while the functional organization of the visual system is well established in the human brain, ANNs only model functional tuning and not their spatial organization. The work that she presented compared ANNs with different architectures and learning rules to brain data, to better understand the architecture and learning rules implemented by the brain. By assigning a 2D position to each unit in each convolutional layer and adding a spatial constraint to the loss function, they found that the patterns of functional organization in the ANN mirrored that of the human visual system. She argued that this approach is key to understanding the constraints that are responsible for architecture and organization of the human brain.

Regardless of the level of abstraction, this GAC highlighted the importance of grounding models of the brain in real neural data. Dr. Nikolaus Kriegeskorte from Columbia University pointed out that in the age of big data and big models, researchers need to make sure that the data reflects true empirical constraints faced by the biological system and that the models capture these constraints. He argued that information about the world and human behavior should be used to set the parameters of the models while the neural data should be used to adjudicate between and validate the proposed models. Only then can researchers build models that can inform our understanding of the brain and produce human-like levels of behavior across multiple tasks and contexts.

Cognition

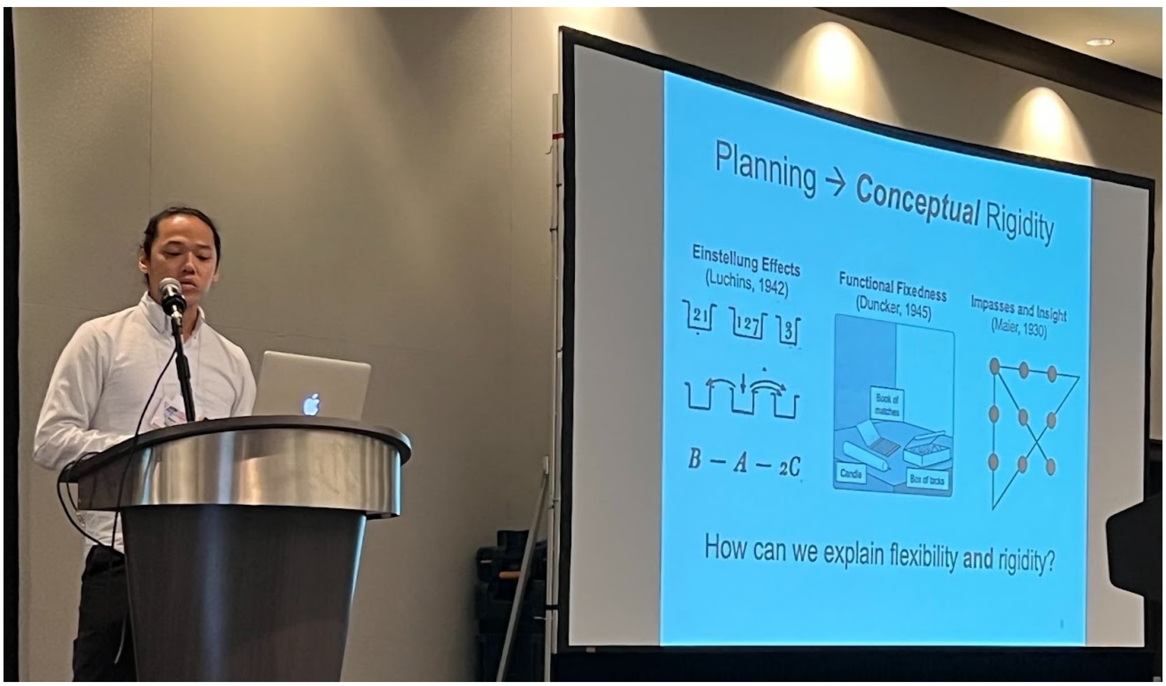
The brain *constructs beliefs about the external world* to be able to predict the best action to take. One school of thought is that the brain simulates the results of actions to determine the best action to take. The second GAC tackled this question: to what extent does the brain simulate the external world?

This question was posed to a large group of panelists, including Dr. Tomer Ullman from Harvard, Dr. Wei Ji Ma from NYU, Dr. Tom Griffiths from Princeton, and Dr. Kelsey Allen from DeepMind. The proponents of mental simulations argued that the brain executes detailed mental simulations akin to a movie being played using a video game physics engine. These intuitive engines can be used for prediction, inference, and learning causality. For example, one study presented by Dr. Kevin Smith from MIT had participants make a decision as to where a stack of blocks would fall. As the number of blocks increased, the time it took for humans to make a decision increased, which correlated with the simulation time of physics engines. Another study using fMRI found representations in the parietal and frontal cortex that match representations from a physics engine. The opposition of the mental simulation group argued that not all decisions require simulations. For example, habits or decisions made with minimal uncertainty are largely automatic decisions. However, most panelists agreed that the brain implements multiple systems for both deliberate and automatic computations. The disagreement centered on the level of detail in the mental simulations.

The panel then moved on to encourage discussion in small groups within the audience. Each discussion was focused on one of three questions. 1) What are the necessary and sufficient components of mental simulation? 2) What is some evidence for mental simulation? 3) What experiment or framework is most promising for studying mental simulation? A majority of researchers agreed that mental simulation requires a step-by-step prediction of the external world. These intermediate steps must be temporally linked, to explain a causal relationship between an action and an outcome. Some examples of evidence for mental simulation were from mental rotation tasks, where the time it took for participants to mentally rotate an object was linearly related to the degree of the rotation. Furthermore, in causal judgment tasks, eye movement data showed that human participants track the predicted trajectory of moving objects, simulating their physical interactions with walls and other agents in their environment. Responses to the third question about how best to study mental simulations were the most heterogeneous, but researchers generally agreed that studying tasks with more complex, naturalistic environments that require mental simulations will be crucial in understanding whether and how the brain computes simulations in everyday life.

Action

Selecting the *most beneficial actions* in a complex environment is traditionally thought to necessitate planning over some abstract representation of the environment. However, the third GAC posed the question of whether or not this is true: is planning actually required for intelligent behavior?



In the earliest artificial models, planning prior to performing the task was necessary, given computational limitations at run time. More recently, in deep reinforcement learning (RL) models, the need for planning was somewhat mitigated by higher computing power and larger data sets. However, the relative importance of planning before or during a task was unresolved. To address this question, Dr. Jessica Hamrick from DeepMind tested the performance of a state-of-the-art deep RL model (MuZero) across multiple tasks and environments. She found that planning prior to performing the task was more important for performance than planning at each decision point during the task. In other words, with the right training, planning is not beneficial for state-of-the-art artificial agents.

The importance of planning compared to non-planning decision making systems was also discussed for biological agents. Dr. Malcolm MacIver from Northwestern University identified the behavioral tradeoffs between the two types of decision making. He noted that at each decision point, animals either make habit-based decisions, which are faster and don't require planning, or plan-based decisions, which are slower and more informed. The type of decision making performed by the animals changes entirely based on the context of the decision. For example, planning is vitally important for situations that are expensive, meaning the result of their decision has fatal ramifications, and are low-N, meaning the animal has not often experienced this situation. Dr. David Redish from the University of Minnesota established a neural basis for multiple decision-making systems. He highlighted work that showed the neural activity recorded from the hippocampus reflected the different types of decision making: theta waves were associated with planning while sharp waves were associated with exploring.

In general, the participants of this panel agreed that planning prior to behaving is crucial for both biological and artificial agents. The presented work suggests that planning can help recover and learn a useful representation of the task and environment. However, the importance of planning during behavior varied between artificial and biological agents. While biological agents face vital decisions that affect their ecological viability as a species, artificial agents are not as harshly punished for making the wrong decision. Thus, the context of the world shapes the importance of planning for intelligent agents.

Learning

The simplest type of error-driven learning requires an agent to *update its knowledge* based on the outcome of its action. Learning is thus an implicit goal, as without it, an agent cannot reach its explicit goal. To help artificial agents learn how to perform an action, researchers typically train them on similar tasks, to learn a mapping between perceptual input and motor action (e.g., training a robotic arm to open a bottle or a legged robot to walk). Yet one major problem highlighted by both Dr. Chelsea Finn from Stanford University and Dr. Deepak Pathak from Carnegie Mellon University is that of generalizing across tasks. Even if the trained tasks are similar to the test, for example testing on opening a bottle with a different cap or walking on different terrain, robots trained with a fixed reward function have trouble performing both variations of the task. Both Dr. Finn and Dr. Pathak modified the learning paradigm, to improve cross-task performance. Dr. Finn found that using prior experience to learn a new task increased the speed of learning and the overall performance on the new task. Dr. Pathak developed a Rapid Motor Adaptation model, which uses continuous domain general simulations to enable the legged robot to walk on a wide range of surfaces with distinct physical properties, such as a steep rocky road or a smooth oily floor.

While generalization is a difficult problem in machines, humans have other limitations, such as a limited memory capacity, redundancy across systems, and the tendency to see structure in randomness. However, Dr. Anne Collins from UC Berkeley argued that such limitations are important features of human intelligence, rather than bugs. She showed that adding a working memory component to a reinforcement learning model created a relationship between the number of objects learned and recognition performance, mirroring human behavior. She argued that limited memory forces the agent to focus on what information is important to learn and that seeing patterns in random noise reflects a rational tendency, as real-world patterns tend to have meaningful structure.

By bringing together researchers from cognitive science, artificial intelligence, and neuroscience, CCN 2022 was able to foster discussions about how to build models of complex human behavior. This year, CCN featured dozens of contributed talks by trainees and early career researchers that we were not able to cover. However, CCN makes all accepted submissions available online here, so we encourage readers to check out their work. By the closing remarks of CCN, it was clear that these types of interdisciplinary discussions are crucial for informing theories about the brain and engineering models that rival human performance.

This report was graciously supported as part of the Tianqiao and Chrissy Chen Institute Science Writers Fellowship. Special thanks to Dr. Thomas Naselaris from the University of Minnesota and Dr. Hendrikje Nienborg from the NIH for their feedback.