Artificial Intelligence and Anthropogeny
Virtual Public Symposium · Friday, March 3, 2023

Co-chairs:
Terry Sejnowski, Salk Institute for Biological Studies & University of California, San Diego
Patricia Churchland, University of California, San Diego

Sponsored by:
Center for Academic Research and Training in Anthropogeny (CARTA)

ABSTRACTS

What Language Models Mean
Blaise Agüera y Arcas, Google

Large language models (LLMs) have now achieved many of the longstanding goals of the quest for generalist AI, and they have done so with large-scale, neurally inspired, attention-enabled, unsupervised machine learning, as opposed to the code- and rule-based approaches that have repeatedly failed over the past half century. While LLMs are still very imperfect (though rapidly improving) in areas like factual grounding, planning, reasoning, safety, memory, and consistency, they do understand concepts, are capable of insight and originality, can problem-solve, and exhibit many faculties we have historically defended vigorously as exceptionally human, such as humor, creativity, and theory of mind. At this point, human responses to the emergence of AI seem to be telling us more about our own psychology, hopes and fears, than about AI itself. However, taking these new AI capacities seriously, and noticing that they all emerge purely from sequence modeling, should cause us to reassess what our own cerebral cortex is doing, and whether we are learning what intelligence, machine or biological, actually is.

Human Languages and their Cognition(s)
Damián Blasi, Harvard University

The emergence of language is routinely regarded as a major (or even the main) evolutionary transition in our species' history. Much less attention and awe has been dispensed to the fact that humans evolved the capacity to successfully create, learn, and use a myriad of different languages which, while similar in some aspects, are radically different in many others. In this presentation, I will argue that these differences have observable consequences for non-linguistic aspects of cognition and behavior. Finally, I will discuss how these effects play out in the design, testing, and deployment of AI, as the linguistic peculiarities of behemoth languages like English are extrapolated to the world's languages.

The Parallel Architecture in Language and Elsewhere
Ray Jackendoff, Tufts University/Massachusetts Institute of Technology

The Parallel Architecture is a theory of the mental representations (or "data structures") involved in the language faculty. These representations are organized in three orthogonal dimensions or levels: phonology (sound structure), syntax (grammatical structure), and semantics (conceptual structure or meaning), correlated with each other through interface links. Words are encoded in all three levels and serve as part of the interface between sound and meaning. In the representation of an entire sentence, the words are spread out across the combinatoriality of the three levels.

An important requirement for a theory of language is that it must offer an account of how we can talk about what we see. It is proposed that conceptual structure in language interfaces with a level of spatial structure – the understanding of physical space – which in turn interfaces with visual, haptic,

and proprioceptive perception, and with the planning of action. A word that denotes a spatial entity (such as cat) links to a representation in spatial structure. Thus, the basic principles of the Parallel Architecture for language can be extended to major aspects of mental function.

The Role of Feedback in the Parallel Architecture of Language
Carmen Amo Alonso & John Doyle, California Institute of Technology

Feedback interconnections are widespread in the brain; yet clear explanations for most of them are currently lacking. We explore current experimental evidence on the relationship between the auditory and motor parts of the brain during speech perception and production, and we propose a simple internal feedback model between the motor system and the auditory system that explains experimental observations. These models provide a plausible explanation for how the structure of language, as described in the Parallel Architecture, is implemented in the brain. Moreover, we provide a plausible account for how the Parallel Architecture of language originates as a result of functional constrains in the sensorimotor system. Furthermore, we compare the brain's implementation of the language capacity with other cognitive capacities, such as vision or motor planning and control, which also have massive internal feedbacks that our new theory explains for the first time. We argue that the conceptual structure of the Parallel Architecture extend beyond the domain of language, and can indeed be applied to the brain more broadly, as well as to other technological tools (particularly those for communication) where both architectures and components are widely known.

The Evolution of Syntax and Pragmatics in a Gradualist Scenario
Eva Wittenberg, Central European University (Vienna)

No account of how people understand language would be complete without an account of pragmatics, the study of how people understand jokes, insinuations, novel metaphors, or subtle nudges — all the meanings beyond the literal meaning that makes our social interactions entertaining, infuriating, creative, or polite, and that pose so much of a headache to developers of artificial systems. But how did language evolve to efficiently relay so much pragmatic trickery? Here, I present a new paper that builds on the idea that grammar evolved gradually, and with it, pragmatics. We argue that the simpler a grammar is, the stronger the reliance on pragmatic inferences for many aspects of meaning, including even basic questions such as who did what to whom. As grammars gradually evolve towards more complex systems, these coarse pragmatic inferences give way to pragmatic processes that are different in character: Syntax, semantics, and the lexicon evolve to contain reliable and systematic triggers for highly structured pragmatic phenomena. Our account thus links a gradualist scenario of the evolution of syntax that triggers distinct qualitative processes in pragmatic reasoning.

Evolution of Birdsong Learning and Human Spoken Language
Erich Jarvis, The Rockefeller University/Howard Hughes Medical Institute

Vocal learning is one of the most critical components of spoken language. It has only evolved several independent times among mammals and birds. Although all vocal learning species are distantly related and have closer relatives that are non-vocal learners, humans and the vocal learning birds have evolved convergent forebrain pathways that control song and speech imitation and production. Here, I will present an overview of the various biological hypothesis of what makes vocal learning and spoken language special, how it evolved, and what differs about the molecular and neural mechanisms compared to other behavioral traits. We find convergent changes in gene regulation in song learning brain pathways in birds and spoken language brain pathways in humans. These genes are enriched for functions in brain connectivity, neural activity, and synaptic plasticity. The specialized regulation is associated with convergent accelerated regulatory regions. To explain these findings, I propose a motor theory of vocal learning origin, in which brain pathways for vocal learning evolved by brain pathway duplication of an ancestral motor learning pathway. The duplicated

pathway uses mostly the same genes, but with divergences in gene regulation via sequence and epigenetic changes, which control divergent connectivity and other specialized functions to rapidly integrate auditory input with vocal motor output.

Linking Communication and Cooperation: Lessons from the Naked Mole-Rat
Alison Barker, Max Planck Institute for Brain Research

Highly organized social groups require well-structured and dynamic communication systems. Naked mole-rats form some of the most rigidly structured social groups in the Animal Kingdom, exhibiting eusociality, a type of highly cooperative social living characterized by a reproductive division of labor with a single breeding female, a queen. Recent work from our group identified a critical role for vocal communication in the organization and maintenance of naked mole-rat social groups. Using machine learning techniques we demonstrated that one vocalization type, the soft chirp, encodes information about individual identity and colony membership. Colony specific vocal dialects can be learned early in life- pups that were cross-fostered acquired the dialect of their adoptive colonies. We also demonstrate that vocal dialects are influenced in part by the presence of the queen. Here, I summarize these findings and highlight our current work investigating how social and vocal complexity evolved in parallel in closely related species throughout the Bathyergidae family of African mole-rats.

Common Sense and AI
Gerd Gigerenzer, Max Planck Institute for Human Development

Common sense is shared knowledge about people and the physical world, enabled by the biological brain. It comprises intuitive psychology, intuitive physics, and intuitive sociality. Unlike deep neural networks, common sense requires only limited experience. Human intelligence has evolved to deal with uncertainty, independent of whether big or small data are available. Complex AI algorithms, in contrast, work best in stable, well-defined situations such as chess and Go, where large amounts of data are available. This stable-world principle helps to understand what statistical algorithms are capable of and distinguish it from commercial hype or techno-religious faith. I introduce the program of psychological AI, which uses psychological heuristics to make algorithms smart. For instance, when predicting the spread of the flu, a situation of uncertainty, the recency heuristic which relies on only one data point can lead to better predictions than Google Flu Trends' big-data algorithms. What we need is a fusion of the adaptive heuristics that embody common sense with the power of machine learning. Reference: Gigerenzer, G. (2022). How to stay smart in a smart world. MIT Press.

Learning by Experiment: Continually Evolving Machines
Pulkit Agrawal, Massachusetts Institute of Technology

Evolution always presented life forms with new challenges -- due to changes in weather, terrain, competition between different organisms, and other reasons. To increase the chance of survival, instead of solely optimizing current performance, it is in an agent's interest to maximize its ability to adapt to changes. Possibly this old evolutionary trait manifests itself in modern humans in their ability to adapt to new tasks and challenges quickly. Even if we consider the lifetime of a human, the ability to adapt is critical. An open question is what enables humans to adapt, a trait that modern AI systems lack. A prominent theory in developmental psychology suggests that "seemingly" frivolous play is a mechanism for infants to experiment to increase their knowledge incrementally. Play prepares infants for future life by laying down the foundation of a high-level experimentation framework to quickly understand how things work in new environments for constructing goal-directed plans. I will discuss how the idea of experimentation can be leveraged to construct robots that improve with experience and solve novel problems presented to them.