

Bridging the Gap Between Neurons and Cognition Through Assemblies of Neurons

Christos H. Papadimitriou

christos@columbia.edu

Columbia University, New York, NY 10027, U.S.A.

Angela D. Friederici

friederici@cbs.mpg.de

Max Planck Institute for Human Cognitive and Brain Sciences, Department of Neuropsychology, D-04303 Leipzig, Germany

During recent decades, our understanding of the brain has advanced dramatically at both the cellular and molecular levels and at the cognitive neurofunctional level; however, a huge gap remains between the microlevel of physiology and the macrolevel of cognition. We propose that computational models based on assemblies of neurons can serve as a blueprint for bridging these two scales. We discuss recently developed computational models of assemblies that have been demonstrated to mediate higher cognitive functions such as the processing of simple sentences, to be realistically realizable by neural activity, and to possess general computational power.

1 Introduction ---

Over the past decades, neuroscientific research at the molecular and cellular levels has dramatically increased our understanding of the physiology and function of neurons, as well as their interconnectivity. At the other end of the scale, cognitive neuroscience has advanced considerably our understanding of the functional neuroanatomy of higher cognitive functions such as memory and language. The gap between the microlevel of physiology and the macrolevel of functional neuroanatomy, however, remains huge. This is due not only to the difference in granularity but also to current methodological difficulties in experimental measurements at the levels in between.

There is another gap separating the two scales: cognitive imaging studies are for the most part carried out on human subjects, while neuron-level neuroscientific experiments typically involve laboratory animals. However, comparative anatomical and electrophysiological studies across different species provide compelling support for the idea that the computations realized in local cortical networks enabling perceptual, motor and cognitive functions are remarkably similar (Balaban et al., 2010; Douglas & Martin,

2004; Markov & Kennedy, 2013; Bullmore & Sporns, 2009). Modeling human cognitive function based on findings from animal studies is not unjustified (Harris, 2005; Friederici & Singer, 2015); see, for example, the direct comparison of event-related brain potentials to processing auditory stimuli between human and nonhuman primates (Milne et al., 2016). It has been argued (Schomers, Garagnani, & Pulvermüller, 2017) that differences in cognitive abilities between humans and animals, such as language, may be attributed mainly to differences in connectivity between distal areas.

The need for bridging the gap between these two established approaches to understanding the brain has been articulated recently by Richard Axel (2018), who states in an interview, “[We need] a logic for the transformation of neuronal activity into thought and action.”

What kind of computational model, mediating between the single-cell level and the cognition level, would fit the bill? Several genres of computational models, operating between the two scales, have been developed over the past decades, and in a recent review (Kriegeskorte & Douglas, 2018, Figure 3) they are categorized into three kinds according to the twin criteria of cognitive fidelity and biological fidelity: (1) models of cognition (Bayesian and graphic models, reinforcement learning models, and cognitive programs, among many others) are good at expressing cognitive functions but lack in biological plausibility; (2) models from computational neuroscience can implement, through biologically faithful model neurons—and their circuits and dynamics—only relatively low-level cognitive functions, such as working memory; and (3), between these two in both axes, deep neural networks are capable of certain aspects of intelligent behavior but have been broadly recognized to violate important biological constraints, such as local sparsity, recurrence, lack of supervision, as well as information propagation from postsynaptic to presynaptic loci (but see Lillicrap, Santoro, Marris, Akerman, & Hinton, 2020, for a recent argument about the biological plausibility of deep nets). We believe that none of these families of models promises to bridge the gap—that is, to explicate “the transformation of neuronal activity into thought and action” (Axel 2018).

2 Cognitive Modeling by Assemblies

D. O. Hebb (1949) proposed that there are persistent and densely interconnected populations of neurons whose near-synchronous firing is coterminous with the subject thinking of an object, concept, or word; he called such populations *assemblies*. Assemblies of neurons were sought by neuroscientists for more than five decades after Hebb’s prediction, until they were observed experimentally in the animal brain (Harris, Csicsvari, Hirase, Dragoi, & Buzsáki, 2003). Novel neuroscientific methods have since been developed (Carillo-Reid, Han Yang Akrouh, & Yuste, 2019) for detecting, measuring, and manipulating assemblies. Results reported in Ison,

Quiroga, and Fried (2015) and elsewhere suggest that in the human medial temporal lobe (MTL), a typical assembly has sparsity roughly 10^{-3} (i.e., an assembly comprises about one in every 1000 excitatory neurons in the specific brain area), and consists of roughly 10^4 neurons, or a small multiple of this number. There is a growing consensus (Harris, 2005; Buszáki, 2010, Buszáki, 2019; Eichenbaum, 2018) that assemblies play an important role in the way the brain works. Since their scale falls squarely between the two extreme scales of single neurons and cognition discussed above, assemblies hold definite promise of bridging the gap.

Indeed, computational models of cognition based on assemblies have been proposed recently. In the 1990s, researchers seeking an explanation for the continuing firing—after the stimulus was no longer present—of neurons in the prefrontal cortex of animals engaged in a working memory task (Amit, Brunel, & Tsodyks, 1994; Amit, Fusi, & Yakovlev, 1997; see also the review by Durstewitz, Seamans, & Sejnowski, 2000) hypothesized that such a neuron must be part of an assembly encoding the stimulus, an assembly that was formed by the stimulus through recurrent excitation and Hebbian plasticity. In two recent papers (Kunze, Peterson, Haueisen, & Knösche, 2017; Kunze, Haueisen, & Knösche, 2019), the analysis of the dynamics of canonical microcircuits of interacting populations of excitatory and inhibitory neurons in cortex reveals three basic modes of these systems, all three corresponding to cognitive behaviors: gating (choosing one among several synaptic input streams), memory (creating a stable assembly), and priming (lowering a perceptual threshold). The authors also demonstrate how such microcircuits can facilitate cognitive functions and, in particular, language understanding (much more on language below).

How are new assemblies created? A picture emerges from the above and other related works. Synaptic input from an extant assembly—perhaps a primitive assembly representing a stimulus at the hippocampus (Quiroga, 2016) or in the olfactory bulb (Franks et al., 2011)—causes a population of neurons in a downstream area to fire. Inhibitory reaction in the downstream area helps select a few of the neurons in the downstream area to keep firing. Next, continuing afferent synaptic input, combined with the recurrent excitation from the firing neurons in the downstream area, results in the firing of a slightly different set of neurons, and then yet another. With continuing excitation from the upstream assembly, eventually a stable set of excitatory neurons will be selected in the downstream area, and these will form the new assembly. The resulting assembly has strong internal synaptic connectivity for two reasons. First, the way its excitatory neurons were selected favors neurons with synapses from within the set; furthermore, the repeated firing increases synaptic strengths through Hebbian plasticity. The new assembly also has strong synaptic connectivity from the parent assembly, again due to Hebbian plasticity: if the parent assembly fires again, the newly formed assembly will follow suit.

Importantly, this process can now be repeated. The new assembly, if excited repeatedly, can create another assembly, in an area further downstream, in the exact same manner. This seems to be an important behavior of assemblies, which can be called *projection*. Projection creates a new “copy” of an assembly, in a different brain area, that will henceforth fire every time the parent assembly does, providing a mechanism whereby memories formed at the hippocampus can be propagated further in the cortex, thus mediating cognition, reasoning, planning, language, and so forth.

In a recent paper, Pokorny et al. (2020) verified through the simulation of a circuit of realistic model spiking neurons with spike timing-dependent plasticity (STDP) that indeed an assembly established in one brain area can create a projection in another brain area and that this new assembly will henceforth behave as described above. Furthermore, the simultaneous presentation of two parent assemblies in the upstream area will modify, through plasticity, the corresponding two assemblies in the downstream area, so they overlap more. Such an increase in overlap, recording the affinity, or *association*, of the two world objects represented by the two assemblies, had been recently noted in experiments in humans (Ison et al., 2015). By presenting the image of a person known to an individual (say, a sibling) at a particular place (say, the pyramids) resulted in a neuron that theretofore only responded to images of the pyramids to also respond to images of the sibling. Presumably this means that two extant assemblies representing the person and the place were modified in response to the presentation of the combined image to increase their overlap.

Besides *projection* and *association*, a third behavior of assemblies, this time related to language processing, has been recently captured by neurorealistic simulations (Müller, Papadimitriou, Maass, & Legenstein, 2020). It is broadly accepted that words such as *cat* are encoded in a largely permanent manner in a subarea of the left MTL containing the medial temporal gyrus (MTG) and usually referred to as *lexicon*. The results in Müller et al. (2020) suggest that the assembly representing the word *cat* in the lexicon can be *bound* to the syntactic category subject in the context of the sentence, “The cat chased the dog.” This binding can be accomplished through a sophisticated variant of projection, involving the creation of a newly formed projected copy of the word *cat* without changing substantially the representation of the word in the lexicon. Indeed, in a recent MRI experiment (Frankland & Greene, 2015) two different subareas of the superior temporal gyrus (STG) were found to exhibit activity in response to the word *cat* in the sentences (1) “The cat chased the dog,” where *cat* is the subject. and (2) “The dog chased the cat,” where *cat* is the object of the sentence. We hypothesize that this activity reflects the binding operation.

Note that the category subject can be encoded in a sentence by different means depending on the language and sentence type. It could be encoded by word position in a sentence (the first noun is the subject in

English or French), or by case marking in a language like German or Russian (the nominative case indicates subject). In a passive sentence, the subject can be marked by the preposition *by* (“by the man” where the thematic role of the actor is indicated by the preposition). As already mentioned, brain data and analysis by Frankland and Greene (2015) suggest that there may be two distinct brain subareas in the STG to which new assemblies are created for the subject and the object (or agent and patient) of a sentence during syntactic processing.

Now let us consider how this binding of words to syntactic roles can be accomplished. Consider two brain regions, one containing the lexicon (stable representations of words) and another area (possibly located at the posterior part of the superior temporal lobe, STL) corresponding to the syntactic role of subject, and suppose that there is synaptic connectivity from the lexicon area to the subject area, *and back*. If the *cat* assembly fires (presumably because the auditory or visual stimulus *cat* was presented) and after the syntactic role of “cat” as subject has been decided (a process discussed below), then synaptic input will excite neurons in the subject area, as in the beginning of the projection operation. But now, because of the backward connectivity, from subject to lexicon, there will be afferent excitation to the “cat” assembly. Eventually the projection in the subject area will stabilize. When this happens—and the simulation experiments in Müller et al. (2020) show that it does happen—the new assembly in the subject area will have strong synaptic connectivity to and from the “cat” assembly in the lexicon area: “cat” has been bound to “subject” through a new assembly created in the subject area. This behavior of assemblies, which can be called *binding* or *reciprocal projection*, was verified in Müller et al. (2020) through a realistic model of spiking neurons.

In another recent paper, Papadimitriou, Vempala, Mitropolsky, Collins, and Maass (2020) proposed that assemblies are capable of an even broader repertoire of operations: the three operations project, associate, and bind that we have seen, plus two more: pattern completion and merge. *Pattern completion* entails an assembly firing in response to a few of its neurons (say, 10% to 20%) firing; this is a consequence of the dense synaptic connectivity within the assembly. Finally, *merge* is an extension of binding: two assemblies, in two different brain areas A and B with synaptic connectivity to a third area C and back, create, through the firing of their parent assemblies, a new assembly in area C, which has strong synaptic connectivity to and from the two original assemblies. Linguists have long believed that human brains must be capable of such a merge operation, whereby two words (or other syntactic elements) are combined to form a more complex syntactic element (Chomsky, 1995). By iterating such merges, arbitrarily large syntactic trees can be formed, capturing the hierarchical structure of language. In fact, it has been verified experimentally (Zaccarella & Friederici, 2015; Zaccarella, Meyer, Makuuchi, & Friederici, 2017) that on completion of a phrase or a sentence, there is neural activity in the posterior portion of Broca’s area, a

brain area believed to be implicated in syntactic analysis; such activity could very well be the result of a merge assembly operation.

At the neurophysiological level, one could ask why these particular five assembly operations—projection, association, pattern completion, binding, and merge—represent realistic behaviors of neurons and synapses. The argument is three-fold. First, there is mathematical proof in a simplified and analytically tractable abstract model of spiking neurons in cortex (Papadimitriou & Vempala, 2019) that these operations can indeed be simulated by stylized neurons and synapses. Second, as noted above, all of these operations have been corroborated through simulations in neurorealistic environments (Müller et al., 2020; Pokorný et al. 2020) and by simulations of the abstract model presented in Papadimitriou et al., (2020). Finally, these operations are consistent with, and appear to provide an explanation for, several experimental results mentioned above—Frankland and Greene (2015), Ison et al. (2015), and Zaccarella and Friederici (2015), among many others.

In terms of computation, these five operations, complemented with three more—*inhibition* and *disinhibition* of an area, *firing* (causing a particular assembly to fire), and *readout* (identifying the assembly that fired in a particular area)—comprise a repertoire of operations on assemblies called the Assembly Calculus. Papadimitriou et al. (2020) proved that the Assembly Calculus is a full-fledged computational framework capable of arbitrary computations, akin to programming languages and the Turing machine (Turing, 1937). Such expressive power is an important desideratum for a framework hypothesized to underlie general cognitive functions. (See section 4 for more about the computational power of the Assembly Calculus.) As we mentioned above, there is evidence (through both mathematical proof and simulations) that the primitives of the Assembly Calculus can be “compiled down” to the level of neurons and synapses, in very much the same way that a program written in a programming language like Python and Matlab can be translated into machine code and executed on a computer. Furthermore, it has been shown that certain simplified cognitive phenomena related to language processing can be expressed as programs in the Assembly Calculus. Overall, we believe that the Assembly Calculus appears quite well positioned to bridge the gap between neurons and cognition.

3 Language Processing through Assemblies

The Assembly Calculus can help explicate the ways in which neuronal activity results in high-level cognitive phenomena. Here we focus on one important high-level cognitive phenomenon: the processing of language.

Why language? First, language is well described in its functional subparts by linguistic theory (Berwick & Chomsky, 2015). Second, there are a number of functional and structural experiments investigating and defining the neuroanatomical basis of language (Friederici, 2011; Fedorenko,

Hsieh, Nieto-Castanon, Whitfield-Gabrieli, & Kanwisher, 2010; Pallier, Devauchelle, & Dehaene, 2011). Third, it is understood that language is grounded on general neurophysiological principles (Friederici & Singer, 2015).

The neural basis of language has been described as a fronto-temporal network in the left hemisphere involving Broca's area in the inferior frontal gyrus (Brodmann areas (BA) 44 and 45), the medial temporal lobe (MTL) and middle temporal gyrus (MTG), and the superior temporal gyrus (STG). These regions are structurally connected by white matter fiber tracts running dorsally and ventrally between frontal and temporal brain regions. These fiber tracts connect the language-relevant brain regions functionally during language processing (see Figure 1A and Friederici, 2011, for a review). For example, the important dorsal fiber called the *arcuate fasciculus* connects the STG with the posterior portion of Broca's area (BA 44), and it is known to have, in humans, larger fiber density in the left hemisphere than in the right (Nucifora, Verma, Melhem, Gur, & Gur, 2005); furthermore, its degree of maturation in children is correlated with accuracy and the speed of sentence processing (Skeide, Brauer, & Friederici, 2016).

In recent years, new insights have been gained into the workings of this network. It is known that soon after the onset of a sentence (read or heard) and once word forms have been identified, neural activity in the MTL signals that the representations of those words contained in the sentence are "looked up" in the lexicon, which is presumed to consist of a large set of assemblies in the MTL (Xu & Sudhof, 2013; Breitenstein et al., 2005) and related assemblies in the MTG (Lau, Phillips, & Poeppel, 2008). The assembly representing each word in the lexicon besides semantic information presumably encodes information about each word related to its use in language, such as its phonological form (*cat*) and its syntactic category (the word *cat* is a noun), which determines the corresponding roles it can play in the sentence (in the case of noun, subject or object, among others). The STG appears to assign the role of each word in the sentence under consideration. In a sentence such as "The dog chased the cat," *dog* is identified as the subject, and *cat* as the object of the sentence, while *chased* is identified as the verb. It has been proposed that in the STG, the word *dog* affects neural activity in different subareas depending on its syntactic function in the two sentences "the dog chased the cat" and "the cat chased the dog" (Frankland & Green, 2015), implying that new representations of the words are created during processing. At this processing stage, words are bound to syntactic roles such as subject, object, and verb. The details of this process at the neurophysiological level need to be investigated in future studies. As a next step, activity in the subareas of Broca's area BA 44 and BA 45 suggests that the buildup of phrases (syntactically licensed parts of the sentence, such as a verb phrase consisting of a verb and a determiner/noun phrase "chased the cat"; see Figure 1B) are supported by BA 44. Finally, representations of the whole sentence, and presumably its meaning, are constructed in the

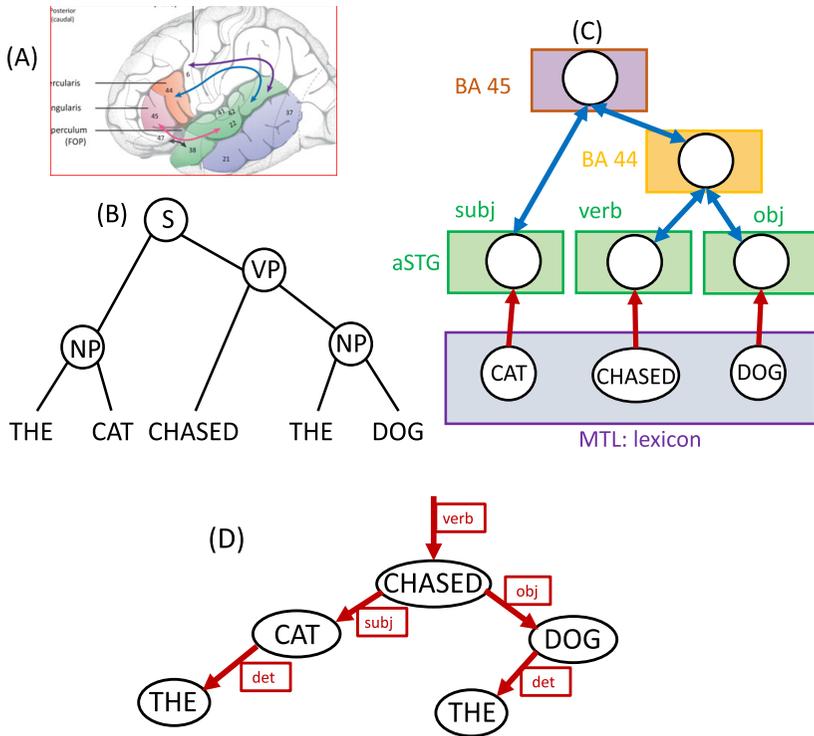


Figure 1: Syntax of language mediated by assemblies. (A) Brain areas implicated in language processing in the left hemisphere. Medial temporal lobe (MTL) is located in the medial part of the temporal cortex behind the middle temporal gyrus (MTG, purple) that is believed to contain the lexicon, representations of all words used in language. The superior temporal gyrus (STG, green) is believed to be implicated in syntactic manipulation of words; the pars opercularis of Broca's area (BA 44, orange) and the pars triangularis (BA 45, magenta) are believed to be implicated in the building of more complex syntactic elements, such as phrases and sentences. Finally, the neural fibers connect the STG and frontal cortex. Two of them (blue, purple) are located dorsally: one of them targets the premotor cortex (BA 6, blue) and the other the posterior portion of Broca's area (BA 44, purple). Two other fiber tracts (red, black) are located ventrally, with the red one targeting the anterior portion of Broca's area (BA 45). BA stands for Brodmann area according to the cytoarchitectonic classification of Brodmann (1909). (B) The syntax tree is a tree-like structure capturing the underlying structure, or syntax, of a sentence. The top node represents the whole sentence (S), while the next lower node represents the verb phrase (VP)—the part of the sentence describing what the subject does or is. Two nodes represent the two noun phrases (NP) corresponding to the subject (left) and the object (right) of the sentence. Each of these noun phrases consists of the determiner *the* and a noun. Identifying the syntax tree of a sentence (or "parsing the

interplay between the temporal cortex and Broca's area under the support of BA 45.

This narrative identifies rather convincingly a very specific brain network for the syntactic analysis of language, which, at least according to the "syntax first" school of thought in language processing, is the first step to comprehension. However, this network cannot be static. The words in the sentence must somehow be "copied" from the lexicon in MTL to the inferior frontal gyrus to determine their syntactic role in the sentence, then bound to the appropriate subarea of STG corresponding to this role. Also, the representations of phrases and sentences in BA 44 and BA 45 must be created through communication with the representations of their subparts and presumably remain in communication with them after the new representations are formed. Such communication between brain regions is possible via the brain's white matter connections: the long-range connections between the language-related regions in the frontal and temporal cortex (Catani, Jones, & Ffytche, 2005; Friederici, 2009; Saur et al., 2008), as well as short-range connections within the frontal cortex (Makuuchi, Bahlmann, Anwander, & Friederici, 2009) and the temporal cortex (Upadhyay et al.

sentence") is considered the first important step toward comprehending the sentence. (C) A proposed dynamic architecture involving assemblies and assembly operations in the left temporal and frontal lobes for performing syntactic analysis of language. The colors of the domains of this architecture (boxes) match the colors of areas in panel A. White circles in these boxes are assemblies in the corresponding areas. Starting from below, assemblies representing the presented (through audio or text) words are identified in the lexicon. Their syntactic category (verb, noun) and their syntactic role (verb, subject, object) are determined first. For verbs, this syntactic role is often readily available in the word's representation; distinguishing between object and subject is subtle and relies on word order (English, French) or case (Russian, German). Once the syntactic roles have been determined, the three assemblies representing, in this example, CAT, CHASED, and DOG, are projected through the bind assembly operation (red arrows), each to the appropriate subarea of the STG, creating three new assemblies there. Next, a merge operation creates an assembly in BA 44 representing the verb phrase. Finally, a second merge operation joins the subject with the verb phrase to create the sentence. These operations create a more rudimentary parse tree than panel B, which encodes the syntactic analysis of the sentence. (D) The Parser in Mitropolsky et al. (in press) implements a different, nongrammatical formalism of syntax called a *dependency graph*. For the same sentence, the Parser would generate the dependency graph shown, in which pairs of words are connected by directed edges (in red), each labeled by the kind of dependency connecting the two words (the labels used here are "verb," "subject," "object," and "determiner"). The root dependency always points to the principal verb of the sentence.

2008). But precisely how is this communication between cell populations in different brain areas established and carried out? The second author of this view predicted (Friederici, 2017, published almost three years before Papadimitriou et al., 2020) “that, for language, there are what I call *mirror cell ensembles* through which two distinct brain regions communicate with each other.”

Assemblies and assembly operations provide precisely the apparatus that is needed here (this apparatus is illustrated in Figure 1C). An assembly representing a word in the MTL can be copied via the projection, or binding, operation to the STG and bound to the appropriate subarea corresponding to its grammatical role in the sentence: subject, verb, or object. The important step of the identification of these grammatical roles, which appears to be happening in the inferior frontal gyrus together with the STG, needs further investigation (see Figure 1C for a brief discussion). An assembly representing the verb phrase can be formed by a merge operation of the verb and the object assemblies in BA 44. Finally, another merge of the subject assembly and the verb phrase assembly will form a new assembly representing the whole sentence. This completes the syntactic analysis of the sentence. Along the way, this process actually builds a “syntactic scaffold” of the sentence, a simplified syntactic tree of the sentence with three leaves and two internal nodes (the structure defined by the blue double arrows in Figure 1C), which can be used as a platform for understanding the precise meaning of the sentence, as well as for, for example, further processing, repetition, and memorization.

In a recent paper, Mitropolsky, Collins, and Papadimitriou (in press) describe a program written in an enhanced version of the Assembly Calculus that successfully parses (i.e., identifies the syntactic structure of) reasonably complex English sentences such as, “This morning, the ferocious dog from the next house briefly chased your cat.” The program, called Parser, works exclusively by the spiking of artificial neurons through the commands of the Assembly Calculus, with two bespoke enhancements: (1) there are commands for inhibiting and disinhibiting neural fibers (synapse bundles connecting different brain areas), instead of whole areas only, and (2) it is assumed that each word assembly in the lexicon is equipped with inhibitory neurons that, upon the firing of the assembly, inhibit or disinhibit the appropriate fibers. This latter capability enables the Parser to correctly identify the syntactic role of words, based exclusively on the word’s part of speech encoded in the (dis)inhibition actions, and the current state of the system. Besides the lexicon, specific brain areas are employed, named, for example, “verb,” “subject,” and “adverb,” where the words in the sentence are projected, while the inhibition or disinhibition of fibers prepares each word for projection to the correct area. The output of the Parser is not a syntactic tree as described in the previous paragraph and Figure 1C, but a different object also capturing the syntactic structure of the sentence, called *dependency*

graph. Obviously, other languages will require different uses of this capability; in Mitropolsky et al. (in press), a toy Russian parser is also described and implemented.

The Parser of Mitropolsky et al. (in press) is an important step toward demonstrating that the Assembly Calculus can bridge the gap between the cognitive and the neuronal level in the domain of language, and especially syntactic analysis. But it does not come close to covering all of language. One important aspect of syntax is *recursion and center embedding*—the ability of sentences to embed other sentences as in, “The dog who chased the cat who scratched the baby is back.” Many consider center embedding to be the hallmark of language and the Parser does not currently handle this feature. A brain area can support many assemblies, and presumably the verb area, for example, can contain assemblies corresponding to the verbs of all sentences in the embedding hierarchy simultaneously; however, the precise implementation of center embedding is left in Mitropolsky et al. (in press) as an open challenge.

4 Conclusion and Open Questions

We have presented some evidence that assemblies of neurons can be a useful conceptual, even mathematical and computational, tool for bridging the forbidding gap between neurons and cognition. But of course, even in the language domain where some progress is being made, we still do not know the answers to most important questions. World languages rely on a stunning diversity of features to encode a word’s grammatical role. How can the same brain circuit adapt to a specific encoding, from among all these very different versions, in the infant brain? Recent brain data indicate that the language-specific cues, and the resulting processing demands, lead to language-specific modulation of the fiber tracts connecting the different brain areas (Goucha, Anwender, & Friederici, 2015). Beyond syntax, how is the meaning of a sentence—its semantics—unraveled in the human brain? And how are sentences combined to form stories and discourses?

The use of the Assembly Calculus for tackling language processing raises a theoretical question. Many decades ago, Noam Chomsky (1956) proposed his influential hierarchy of languages for the purpose of characterizing natural language processing problems: regular, context-free, context-sensitive, and recursively enumerable languages. Of these, the last is far too strong, containing uncomputable problems. The question arises,¹ Is the Assembly Calculus computationally powerful enough to encompass the first three levels of the Chomsky hierarchy? The Assembly Calculus is known to be capable of simulating arbitrary computations with $O(s^{-1})$ space (see the SI

¹Many thanks to an anonymous reviewer for bringing this up.

section of Papadimitriou et al. 2020), where s is the sparsity of assemblies believed to be about 10^{-3} . Therefore, what is known about the complexity of these three levels (Papadimitriou, 1995) implies that regular, context-free, and context-sensitive features of language can be handled through the Assembly Calculus for sentences containing at least a few dozen words.

At a more general level, let us revisit briefly the hypothesis that operations on assemblies underlie cognitive processes. According to all evidence, it is highly unlikely that the brain works in the pristine mathematical, if highly random, manner of the assembly operations, as described in Papadimitriou et al. (2020) and sketched here. The interesting question may rather be, Are these operations useful abstractions of the actual neural processes underlying cognition, including the processing of language? Further work is needed to pursue this question.

Finally, can assembly-based mechanisms such as the Assembly Calculus be demonstrated through simulations to be capable of implementing further cognitive functions beyond language (a domain in which very partial progress has already been made)? Two cognitive tasks that come to mind as interesting candidates are planning and reasoning. Extending the methodology developed for handling language to these and other realms of cognition seems both interesting and challenging.

Acknowledgments

We are grateful to Thomas R. Knösche for valuable feedback on a draft of this view and to Stephen Jose Hanson for a number of constructive comments. The research of C. H. P. was supported by NSF Awards CCF1763970 and CCF1910700, by a research contract with Softbank, and by a grant from Columbia's Center for AI Technology. The research of A. F. was supported by the Deutsche Forschungsgemeinschaft, grants FR519/22-1 within SPP2041.

References

- Amit, D. J., Brunel, N., & Tsodyks, M. V. (1994). Correlations of cortical Hebbian reverberations: Theory versus experiment. *Journal of Neuroscience*, *14*, 6435–6445. 10.1523/JNEUROSCI.14-11-06435.1994, PubMed: 7965048
- Amit, D. J., Fusi, S., & Yakovlev, V. (1997). Paradigmatic working memory (attractor) cell in IT cortex. *Neural Computation*, *9*, 1071–1092. 10.1162/neco.1997.9.5.1071, PubMed: 9188192
- Axel, R. (2018). Richard Axel. *Neuron*, *99*(6), 1110–1112.
- Balaban, E., Edelman, S., Grillner, S., Grodzinski, U., Jarvis, E. D., Kaas, J. H., . . . Pipa, G., (2010). Evolution of dynamic coordination. In C. Von der Malsburg, W. A. Phelps, & W. Singer (Eds.), *Dynamic coordination in the brain. From neurons to mind* (pp. 59–82). Cambridge, MA: MIT Press and FIAS.

- Berwick, R. C., & Chomsky, N. (2015). *Why only us: Language and evolution*. Cambridge, MA: MIT Press.
- Breitenstein, C., Jansen, A., Deppe, M., Foerster, A. F., Sommer, J., Wolbers, T., & Knecht, S. (2005). Hippocampus activity differentiates good from poor learners of a novel lexicon. *NeuroImage*, 25(3), 958–968. 10.1016/j.neuroimage.2004.12.019, PubMed: 15808996
- Brodman, K. (1909). Beiträge zur histologischen Lokalisation der Grosshirnrinde: VI. Die Cortexgliederung des Menschen. *Journal für Psychologie und Neurologie*, 10, 231–246.
- Bullmore, E., & Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10, 186–198. 10.1038/nrn2575, PubMed: 19190637
- Buzsáki, G. (2010). Neural syntax: Cell assemblies, synapsembles, and readers. *Neuron*, 68, 362–385.
- Buzsáki, G. (2019). *The brain from inside out*. Oxford: Oxford University Press.
- Carrillo-Reid, L., Han S., Yang W., Akrouh, A., & Yuste, R. (2019). Controlling visually guided behavior by holographic recalling of cortical ensembles. *Cell*, 178(2), 447–457.
- Catani, M., Jones, D. K., & Ffytche, D.H. (2005). Perisylvian language networks of the human brain. *Annals of Neurology*, 57(1), 8–16. 10.1002/ana.20319, PubMed: 15597383
- Chomsky, N. (1956). Three models for the description of language. *IRE Trans. Inform. Theory*, 2, 113–124. 10.1109/TIT.1956.1056813
- Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.
- Douglas, R. J., & Martin, K. A. C. (2004). Neuronal circuits of the neocortex. *Annual Reviews Neuroscience*, 27, 419–451. 10.1146/annurev.neuro.27.070203.144152
- Durstewitz, D., Seamans, J. K., & Sejnowski, T. J. (2000). Neurocomputational models of working memory. *Nature Neuroscience*, 3(11), 1184–1191. 10.1038/81460, PubMed: 11127836
- Eichenbaum, H. (2018). Barlow versus Hebb: When is it time to abandon the notion of feature detectors and adopt the cell assembly as the unit of cognition? *Neuroscience Letters*, 680, 88–93. 10.1016/j.neulet.2017.04.006, PubMed: 28389238
- Fedorenko, E., Hsieh, P. J., Nieto-Castanon, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New method for fMRI investigations of language: Defining ROIs functionally in individual subjects. *Journal of Neurophysiology*, 104(2), 1177–1194. 10.1152/jn.00032.2010, PubMed: 20410363
- Frankland, S. M., & Greene, J. D. (2015). An architecture for encoding sentence meaning in left mid-superior temporal cortex. In *Proceedings of the National Academy of Sciences of the United States of America*, 112(37), 11732–11737. 10.1073/pnas.1421236112
- Franks, K. M., Russo, M. J., Sosulski, D. L., Mulligan, A. A., Siegelbaum, S. A., & Axel, R. (2011). Recurrent circuitry dynamically shapes the activation of piriform cortex. *Neuron* 72, 49–56 (2011). 10.1016/j.neuron.2011.08.020, PubMed: 21982368
- Friederici, A. D. (2009). Pathways to language: Fiber tracts in the human brain. *Trends in Cognitive Sciences*, 13(4), 175–181. 10.1016/j.tics.2009.01.001, PubMed: 19223226

- Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews*, *91*, 1357–1392. 10.1152/physrev.00006.2011, PubMed: 22013214
- Friederici, A. D. (2017). *Language in our brain. The origins of a uniquely human capacity*. Cambridge, MA: MIT Press.
- Friederici, A. D., & Singer, W. (2015). Grounding language processing on basic neurophysiological principles. *Trends in Cognitive Sciences*, *19*(6), 329–338. 10.1016/j.tics.2015.03.012, PubMed: 25890885
- Goucha, T., Anwander, A., & Friederici, A. D. (2015). How language shapes the brain: Cross-linguistic differences in structural connectivity. Poster presented at the 45th Annual Meeting of the Society for Neuroscience, Chicago.
- Harris, K. D. (2005). Neural signatures of cell assembly organization. *Nature Reviews Neuroscience*, *6*(5), 399–407. 10.1038/nrn1669, PubMed: 15861182
- Harris, K., Csicsvari, J., Hirase, Dragoi, G., & Buzsáki, G. (2003). Organization of cell assemblies in the hippocampus. *Nature*, *424*, 552–556. 10.1038/nature01834, PubMed: 12891358
- Hebb, D. O. (1949). *The organisation of behavior*. New York: Wiley.
- Ison, M. J., Quiroga, Q. R., & Fried, I. (2015). Rapid encoding of new memories by individual neurons in the human brain. *Neuron*, *87*(1), 220–230. 10.1016/j.neuron.2015.06.016, PubMed: 26139375
- Kriegeskorte, N., & Douglas, P. K. (2018). Cognitive computational neuroscience. *Nature Neuroscience*, *21*, 1148–1160. 10.1038/s41593-018-0210-5, PubMed: 30127428
- Kunze, T., Hauelsen, J., & Knösche, T. R. (2019). Emergence of cognitive priming and structure building from the hierarchical interaction of canonical microcircuit models. *Biological Cybernetics*, *113*, 273–291. 10.1007/s00422-019-00792-y, PubMed: 30767085
- Kunze T., Peterson A. D. H., Hauelsen J., & Knösche T. R. (2017). A model of individualized canonical microcircuits supporting cognitive operations. *PLOS One*, *12*, e0188003. 10.1371/journal.pone.0188003, PubMed: 29200435
- Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (De)constructing the N400. *Nature Reviews Neuroscience*, *9* (12), 920–933. 10.1038/nrn2532, PubMed: 19020511
- Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., & Hinton, G. (2020). Back-propagation and the Brain. *Nature Reviews Neuroscience*, *21*, 335–346. 10.1038/s41583-020-0277-3, PubMed: 32303713
- Makuuchi, M., Bahlmann, J., Anwander, A., & Friederici, A. D. (2009). Segregating the core computational faculty of human language from working memory. In *Proceedings of the National Academy of Sciences of the USA*, *106*, 8362–8367. 10.1073/pnas.0810928106, PubMed: 19416819
- Markov, N. T., & Kennedy, H. (2013). The importance of being hierarchical. *Current Opinion in Neurobiology*, *23*, 187–194. 10.1016/j.conb.2012.12.008, PubMed: 23339864
- Milne, A. E., Mueller, J. L., Männel, C., Attaheri, A., Friederici, A. D., & Petkov, C. I. (2016). Evolutionary origins of non-adjacent sequence processing in primate brain potentials. *Scientific Reports*, *6*: 36259. 10.1038/srep36259, PubMed: 27827366

- Mitropolsky, D., Collins, M., & Papadimitriou, C. H. (in press). A biologically plausible parser. *Transactions of the Association for Computational Linguistics*. arXiv: abs/2108.02189.
- Müller M. G., Papadimitriou, C. H., Maass, W., & Legenstein, R. (2020). A model for structured information representation in neural networks of the brain. *eNeuro*, 7(3). 10.1523/ENEURO.0533-19.20207.
- Nucifora, P. G., Verma, R., Melhem, E. R., Gur, R. E., & Gur, R. C. (2005). Leftward asymmetry in relative fiber density of the arcuate fasciculus. *NeuroReport*, 16(8), 791–794. 10.1097/00001756-200505310-00002
- Pallier, C., Devauchelle, A.-D., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. In *Proceedings of the National Academy of Sciences of the United States of America*, 108(6), 2522–2527. 10.1073/pnas.1018711108
- Papadimitriou, C. H. (1995). *Computational complexity*. Reading, MA: Addison-Wesley.
- Papadimitriou, C. H., & Vempala, S. S. (2019). Random projection in the brain and computation with assemblies of neurons. *10th Innovations in Theoretical Computer Science*. Wadernan, Germany: Leibniz International Proceedings in Informatics, Schloss Dagstuhl.
- Papadimitriou, C. H., Vempala, S. S., Mitropolsky D., Collins, M. J., & Maass, W. (2020). In *Proceedings of the National Academy of Sciences of the United States of America*, 117(25), 14464–14472.
- Pokorny, C., Ison, M. J., Rao, A., Legenstein, R., Papadimitriou, C. H., & Maass, W. (2020). STDP forms associations between memory traces in networks of spiking neurons. *Cerebral Cortex*, 30(3), 952–968. 10.1093/cercor/bhz140, PubMed: 31403679
- Quiroga, R. Q. (2016). Neuronal codes for visual perception and memory. *Neuropsychologia*, 83, 227–241. 10.1016/j.neuropsychologia.2015.12.016
- Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.-S., . . . Weiller, C. (2008). Ventral and dorsal pathways for language. In *Proceedings of the National Academy of Sciences*, 105(46), 18035–18040. 10.1073/pnas.0805234105
- Schomers, M. R., Garagnani, M., & Pulvermüller, F. (2017). Neurocomputational consequences of evolutionary connectivity changes in perisylvian language cortex. *Journal of Neuroscience*, 37, 3045–3055. 10.1523/JNEUROSCI.2693-16.2017, PubMed: 28193685
- Skeide, M. A., Brauer, J., Friederici, A. D. (2016). Brain functional and structural predictors of language performance. *Cerebral Cortex*, 26(5), 2127–2139. 10.1093/cercor/bhv042, PubMed: 25770126
- Turing, A. M. (1937). On computable numbers, with an application to the Entscheidungsproblem. In *Proceedings of the London Mathematical Society*, 42, 230–265. 10.1112/plms/s2-42.1.230
- Upadhyay, J., Silver, A., Knaus, T. A., Lindgren, K. A., Ducros, M., Kim, D.-S., & Tager-Flusberg, H. (2008). Effective and structural connectivity in the human auditory cortex. *Journal of Neuroscience*, 28(13), 3341–3349. 10.1523/JNEUROSCI.4434-07.2008, PubMed: 18367601
- Xu, W., & Sudhof, T. C. (2013). A neural circuit for memory specificity and generalization. *Science*, 339(6125), 1290–1295. 10.1126/science.1229534, PubMed: 23493706

- Zaccarella, E., & Friederici, A. D. (2015). Merge in the human brain: a sub-region based functional investigation in the left pars opercularis. *Frontiers in Psychology*, 6, 1818. 10.3389/fpsyg.2015.01818, PubMed: 26640453
- Zaccarella, E., Meyer, L., Makuuchi, M., & Friederici, A. D. (2017). Building by syntax: The neural basis of minimal linguistic structures. *Cerebral Cortex*, 27(1), 411–421. 26464476

Received April 26, 2021; accepted September 2, 2021.