A Calculus for Brain Computation

Christos H. Papadimitriou
Columbia University

Santosh S. Vempala
Georgie Tech

Daniel Mitropolsky
Columbia University

Michael Collins
Columbia University

Wolfgang Maass
Graz U. of Technology

Larry F. Abbott
Columbia University

Abstract

Do brains compute? How do brains learn? How are intelligence and language achieved in the human brain? In this pursuit, we develop a formal calculus and associated programming language for brain computation, based on the assembly hypothesis, first proposed by Hebb: the basic unit of memory and computation in the brain is an assembly, a sparse distribution over neurons. We show that assemblies can be realized efficiently and neuroplausibly by using random projection, inhibition, and plasticity. Repeated applications of this RP&C primitive (random projection and cap) lead to (1) stable assembly creation through projection; (2) association and pattern completion; and finally (3) merge, where two assemblies form a higher-level assembly and, eventually, hierarchies. Further, these operations are composable, allowing the creation of stable computational circuits and structures. We argue that this functionality, in the presence of merge in particular, might underlie language and syntax in humans.

Keywords: assemblies; ensembles; neural computation; neural syntax; language

Introduction

How does the mind emerge from the brain? How do molecules, neurons, and synapses orchestrate cognition, behavior, intelligence, reasoning, language? The remarkable and accelerating progress in neuroscience, both experimental and theoretical-computational, does not seem to bring us closer to the answer, and a new conceptual framework seems to be needed. As Richard Axel recently put it (A, 2018) *"We do not have a logic for the transformation of neural activity into thought and action. I view discerning [this] logic as the most important future direction of neuroscience"*.

What kind of formal system, abstracting the realities of neural activity, would qualify as the sought "logic"? *This is the question pursued in this work.*

We propose a formal computational model of the brain based on *assemblies of neurons,* occupying a position intermediate, in terms of detail and granularity, between the level of individual neurons and synapses, and the level of the whole brain models useful in cognitive science, e.g. (Lake, Salakhutdinov, & Tenenbaum, n.d.). We call this system the *Assembly Calculus.* The basic elementary object of our system is the *assembly of excitatory neurons*. Assemblies were first hypothesized seven decades ago by Donald O. Hebb (Hebb, 1949) as densely interconnected sets of neurons whose loosely synchronized firing in a pattern is coterminous with the subject thinking of a particular concept or idea. Assembly-like formations have been sought by researchers during the decades following Hebb's prediction, see for example (Abeles, 1991), until they were clearly identified more than a decade ago through calcium imaging (Harris, 2005; Buzsaki, 2010). More recently, assemblies (sometimes called *ensembles*) and their dynamic behavior have been studied extensively in the animal brain, see for example (Miller, Ayzenshtat, Carrillo-Reid, & Yuste, 2014).

Our calculus outfits assemblies with certain *operations* that create new assemblies and/or modify existing ones: *project, reciprocal-project, associate, merge,* and a few others. These operations reflect known properties of assemblies observed in experiments, and they can be shown, either analytically or through simulations (almost always both), to result from the activity of neurons and synapses. In other words, the high-level operations of this system can be "compiled down" to the world of neurons and synapses — a fact reminiscent of the way in which high-level programming languages are translated into machine code.

Mathematical Model

Our mathematical and simulation results employ a simplified and analytically tractable, yet plausible, model of neurons and synapses, which we believe may be of broader interest and use. We assume a finite number of brain areas denoted A, B, C, etc., intended to stand for an anatomically and functionally meaningful partition of the cortex. Each area contains n excitatory neuronswith *random* recurrent connections. By this we mean that each ordered pair of neurons in an area has the same small probability p of being connected by a synapse, independently of what happens to other pairs — this is a well studied framework usually referred to as $G_{n,p}$ (Erdős & Renyi, 1960). In addition, for certain ordered pairs of areas, say (A, B), there are random *afferent* synaptic interconnections from A to B; that is, for every neuron in A and every neuron in B there is a chance p that they are connected by a synapse.[1] Our analytical results establishing the effectiveness of the various operations contain the clause "with high probability," where the event space is implicitly the underlying

---

[1]Several experimental results (Song, Sjöström, Reigl, Nelson, & Chklovskii, 2005; Guzman, J Schlö gl, Frotscher, & Jonas, 2016) suggest deviations of the synaptic connectivity of the animal brain from the uniformity of $G_{n,p}$. Such deviations mostly support our results, albeit in ways we are not discussing here.

random graph. We assume that all cells in an assembly x belong to the same brain area, denoted area(x).

We assume *multiplicative Hebbian plasticity:* if at a time step neuron i fires and at the next time step neuron j fires, and there is a synapse from i to j, the weight of this synapse is multiplied by $(1+\beta)$, where $\beta > 0$ is the final parameter of our model (along with n, p, and k). Larger values of the plasticity coefficient β result in the operations converging faster, and render many of our proofs simpler. The plasticity process is complemented by *homeostasis,* namely, the incoming synaptic weights at a neuron are normalized over a larger time scale to maintain their sum constant. Finally, we model *inhibition and excitatory–inhibitory balance* by postulating that neurons fire in discrete time steps, and *at any time only a fixed number k of the excitatory neurons in any area fire.*

The four basic parameters of our model are these: n (the number of excitatory neurons in an area, and the basic parameter of our model); p (the probability of recurrent and afferent synaptic connectivity); k (the maximum number of firing neurons in any area); and the plasticity coefficient β. Typical values of these parameters in our simulations are $n = 10^7, p = 10^{-3}, k = 10^4, \beta = 0.05$. We sometimes assume that k is (a small multiple of) the square root of n; this extra assumption seems compatible with experimental data, and yields certain interesting further insights.

Remark: Our model is *generic* in the sense that it is not assumed that circuits specific to various tasks are already in place. Its functionality — the needed apparatus for each task such as implementing an assembly operation — emerges from the randomness of the network and the selection of the k neurons with highest synaptic input as an almost certain consequence of certain simple steps — such as the repeated firing of an assembly.

Assembly Operations

1. Projection How do assemblies in the association cortex come about? It has been hypothesized (see e.g. (Quiroga, 2016)) that an assembly imprinting, for example, a familiar face in a subject's medial temporal lobe (MTL) is created by the *projection* of a neuronal population, perhaps in the inferotemporal cortex (IT), encoding this face as a whole object. By *projection* of an assembly x in area A we mean the creation of a new assembly y in a downstream area B that can be thought of as a "copy" of x, and such that y will henceforth fire every time x fires — an operation denoted project(x, B, y). The process was vividly predicted in the discussion section of (Franks et al., 2011) for the case in which A is the olfactory bulb and B the piriform cortex, the process entails the repeated firing of x. Once x has fired once, synaptic connectivity from area A to B excites many neurons in area B. Inhibition will soon limit the firing in area B to a smaller set of neurons, let us call it y_1, consisting in our framework of k neurons. Now, the simultaneous firing of x and y_1 creates a stronger excitation in area B (one extra reason for this is *plasticity*, which has already strengthened the connections from x

to y_1), and as a result a new set of k neurons from area B will be selected to fire, call it y_2. If x continues firing, a sequence $\{y_t\}$ of sets of neurons of size k in area B will be created. For a large range of parameters and for high enough plasticity, we have proved this process to converge exponentially fast, with high probability, to create an assembly y, the result of the projection, and the trade-off between the speed of convergence and the strength of plasticity has been calculated almost exactly (Papadimitriou & Vempala, 2019). Assembly projection has also been demonstrated through neuromorphic simulations (Pokorny et al., 2017; Legenstein, Papadimitriou, Vempala, & Maass, 2016). Once the project(x, B, y) operation has taken place, we write $B =$ area(y) and $x =$ parent(y).

Hebb (Hebb, 1949) hypothesized that assemblies are *densely interconnected* — that is, the chance that two neurons have a synaptic connection is significantly larger when they belong to the same assembly — and our analysis and simulations of the project operation verify this hypothesis.
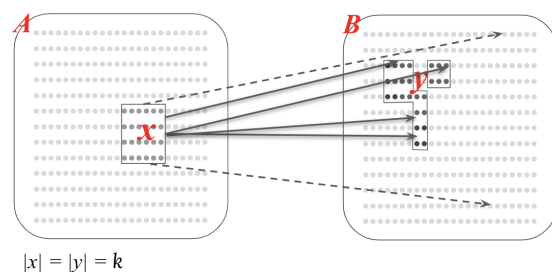

$$|x| = |y| = k$$

Figure 1: Schematic of project(x, B, y).

2. Association In a recent experiment (Ison, Quiroga, & Fried, 2015), electrocorticography (eCoG) recordings of human subjects revealed that a neuron in a subject's MTL consistently responding to a stimulus consisting of the image of a particular familiar place — such as the Pyramids — starts to also respond to the image of a particular familiar person — say, the subject's sibling — once a combined image has been shown of this person in that place. A compelling parsimonious explanation of this and many similar results is that two assemblies imprinting two different entities adapt to the cooccurrence, or other observed affinity, of the entities they imprint by increasing their overlap, with cells from each migrating to the other while other cells leave the assemblies to maintain its size to k; we say that the two assemblies are *associating* with one another. The association of two assemblies x and y *in the same brain area* is denoted associate(x, y), with the common area area(x) = area(y) implicit. We have proved analytically that the simultaneous sustained firing of the two parents of x and y does effect such increase in overlap (Papadimitriou & Vempala, 2019) (this was also demonstrated through neurorealistic simulations (Pokorny et al., 2017)); and also that,

if two associated assemblies are projected to a downstream area, the extent of their overlap is preserved (Papadimitriou & Vempala, 2019). The association operation evokes a conception of a brain area as the arena of complex overlap patterns between the area's assemblies; for an analytical treatment of this view see (Anari et al., 2018).

3. Pattern completion Pattern completion is the firing of the whole assembly in response to the firing of a small number of its cells, which happens with a certain a priori probability. This phenomenon has been demonstrated both analytically and in simulations, and has been observed in vivo (Miller et al., 2014). We believe that association and pattern completion open up fascinating possibilities for a genre of probabilistic computation through assemblies, a research direction which must be further pursued.

4. Merge The most sophisticated, and demanding both analytically and in terms of brain resources, operation in the repertoire of assembly calculus is $\mathrm{merge}(x,y,A,z)$, the creation of a new assembly z in area A that has strong two-way synaptic connectivity with two other assemblies x and y, in two different brain areas. Our algorithm for implementing merge is by far the most complex in this work, as it requires the coordination of *five* different brain areas with ample connectivity between them, as well as stronger plasticity. This is consistent with the hypothesis that merge is an operation particular to humans, and that it has possibly required the development of new heights of neural and synaptic capabilities in the human brain, such as the lateralization and size differentiation of the left arcuate fasciculus (Vernooij et al., 2017). Finally, a simpler operation with similar yet milder complexity is $\mathrm{reciprocal.project}(x,A,y)$: It is an extension of $\mathrm{project}(x,A,y)$, with the additional functionality that the resulting y has strong *backward* synaptic connectivity to x. It has been shown to be implementable through detailed simulations of realistic networks of neurons (Legenstein et al., 2016).

The RP&C primitive It can be said that assembly operations, as described here, are powered exclusively by two forces known to be crucial for life more generally: *randomness* and *selection*. No special-purpose neural circuits are required to be in place; what is needed is *random synaptic connectivity* between, and recurrently within, populations, and selection through inhibition, in each area, of the k out of n cells currently receiving highest synaptic input. All computation consists of applications of this operator, which we call *random projection and cap (RC&P)*. We believe that RP&C is an important primitive of neural computation.

Control operations and computational power

Besides the operations $\mathrm{project}$, $\mathrm{associate}$, $\mathrm{pattern\_complete}$, $\mathrm{reciprocal.project}$ and merge, the Assembly Calculus contains certain basic operations: $\mathrm{read}()$ returns the set of all areas in which an assembly has just fired, see (Buzsaki, 2010) for an argument that, for assemblies

to be functionally useful, readout mechanisms must be in place. Also, we further assume that an assembly x can be explicitly caused to fire by the operation $\mathrm{fire}(x)$, and that the excitatory cells in any area A can be explicitly inhibited and disinhibited by appropriate commands $\mathrm{inhibit}(A)$ and $\mathrm{disinhibit}(A)$. It can be shown that the Assembly Calculus, augmented with these control operations and under reasonable assumptions, is capable of simulating arbitrary $O(\sqrt{n})$ space-bounded computations. Significantly, the simulation relies crucially on the use of the merge operation. This is a remarkable computational capability, consistent with the hypothesis, articulated next, that the Assembly Calculus may underlie the more sophisticated functions of the human brain.

The assembly hypothesis

The Assembly Calculus is a formal system with a repertoire of rather sophisticated operations, where each of these operations (with the exception of the control primitives, which however are arguably also quite plausible) can be ultimately reduced to the firing of randomly connected populations of excitatory neurons with inhibition and Hebbian plasticity. The ensuing computational power of the assembly calculus may embolden one to hypothesize that this computational system — *or something far more complex and detailed, which however can be usefully abstracted this way* — may underlie advanced cognitive functions, especially those of the human brain, such as reasoning, planning, and language.

Language

Language is arguably the crowning functionality of the human brain. Linguists had long predicted that the human brain is capable of the Merge operation, that is, combining, in a particularly strong sense, two separate syntactic entities to create a new entity representing this particular combination (Berwick & Chomsky, 2016; Hauser, Chomsky, & Fitch, 2002), and that this ability is *recursive* in that the combined entity can in turn be combined with others. This is a crucial step in the creation of the *trees* that seem necessary for the syntactic processing of language, but also for hierarchical thinking more generally (e.g., deduction, discourse, planning, storytelling, etc.). The Assembly Calculus operations $\mathrm{project}$, $\mathrm{reciprocal.project}$, and merge provide the basis for a possible implementation of such hierarchies. In particular, we speculate in Figure 2 with one plausible implementation of the *generation of a sentence* such as *"boy kicks ball,"* in a manner that is compatible with recent experimental results on language in the brain, see e.g. (Frankland & Greene, 2015; Ding, Melloni, Zhang, Tian, & Poeppel, 2016; Friederici, 2018). First, the words corresponding to the verb, subject and object of the intended sentence are identified in the lexicon, presumably residing in the medial temporal lobe (MTL), and are reciprocally projected to the subregions of Wernicke's area (as predicted in (Frankland & Greene, 2015)) corresponding to these three syntactic categories by $\mathrm{reciprocal.project}$ assem-

bly operations (marked (1) in the figure). Such *binding* was recently shown to be neuroplausibly implementable (Legenstein et al., 2016). Following this step, a `merge` operation utilizes the *left arcuate fasciculus* fiber connecting the areas of Wernicke and Broca to create a "verb phrase" or VP assembly by combining the verb and the object, see (2) in the figure. Finally, the subject is `mergeed` with VP to form the sentence S (see (3) in the figure). After this bottom-up generation process, the constituents of the resulting sentence S can next be *articulated* by a top-down process, in an order that reflects the grammar of the particular language, to eventually reach the lexicon assemblies and activate the associated speech processes. Even though this account disregards a host of linguistic subtleties, it does illustrate the relevance of assembly operation to language in the brain.
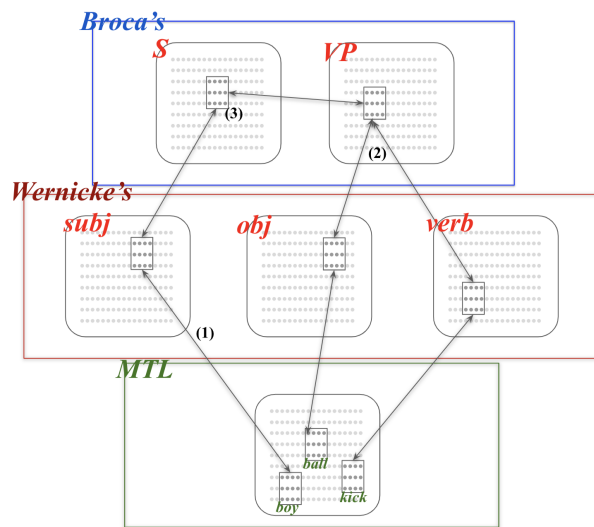

Figure 2: A potential architecture for syntax in the brain.

References

A, Q. . (2018). Richard axel. *Neuron*, *99*.

Abeles, M. (1991). *Corticonics: Neural circuits of the cerebral cortex*. Cambridge University Press.

Anari, N., Daskalakis, C., Maass, W., Papadimitriou, C. H., Saberi, A., & Vempala, S. (2018). Smoothed analysis of discrete tensor decomposition and assemblies of neurons..

Berwick, R. C., & Chomsky, N. (2016). *Why only us: Language and evolution*. MIT Press.

Buzsaki, G. (2010). Neural syntax: cell assemblies, synapsembles, and readers. *Neuron*, *68*(3).

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature neuroscience*, *19*(1).

Erdős, P., & Renyi, A. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hungary. Acad. Sci.*, *5*.

Frankland, S. M., & Greene, J. D. (2015). An architecture for encoding sentence meaning in left mid-superior temporal cortex. *Proceedings of the National Academy of Sciences*, *112*(37).

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*(1).

Friedericci, A. (2018). *Language in our brain*. MIT Press.

Guzman, S., J Schlö gl, A., Frotscher, M., & Jonas, P. (2016). Synaptic mechanisms of pattern completion in the hippocampal ca3 network. *Science*, *353*(6304).

Harris, K. D. (2005). Neural signatures of cell assembly organization. *Nature Reviews Neuroscience*, *6*(5).

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *science*, *298*(5598).

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory*. Wiley, New York.

Ison, M. J., Quiroga, R. Q., & Fried, I. (2015). Rapid encoding of new memories by individual neurons in the human brain. *Neuron*, *87*(1).

Lake, B., Salakhutdinov, R., & Tenenbaum, J. (n.d.). Human-level concept learning through probabilistic program induction. *Science*, *350*.

Legenstein, R., Papadimitriou, C. H., Vempala, S., & Maass, W. (2016). Assembly pointers for variable binding in networks of spiking neurons. *arXiv preprint arXiv:1611.03698*.

Miller, J.-e. K., Ayzenshtat, I., Carrillo-Reid, L., & Yuste, R. (2014). Visual stimuli recruit intrinsically generated cortical ensembles. *Proceedings of the National Academy of Sciences*, *111*(38).

Papadimitriou, C. H., & Vempala, S. S. (2019). Random projection in the brain and computation with assemblies of neurons. In *10th innovations in theoretical computer science conference, ITCS 2019, january 10-12, 2019, san diego, california, USA*.

Pokorny, C., Ison, M. J., Rao, A., Legenstein, R., Papadimitriou, C., & Maass, W. (2017). Associations between memory traces emerge in a generic neural circuit model through STDP. *bioRxiv:188938*.

Quiroga, R. Q. (2016). Neuronal codes for visual perception and memory. *Neuropsychologia*, *83*.

Song, S., Sjöström, P. J., Reigl, M., Nelson, S., & Chklovskii, D. B. (2005). Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biology*, *3*(3).

Vernooij, M., Smits, M., Wielopolski, P., Houston, G., Krestin, G., & van der Lugt, A. (2017). Fiber density asymmetry of the arcuate fasciculus in relation to functional hemispheric language lateralization in both right- and left-handed healthy subjects: A combined fmri and dti study. *NeuroImage*, *35*(3).