The Assembly Hypothesis: Emergent Computation and Learning in a rigorous model of the Brain

Santosh Vempala

Algorithms and Randomness Center, Georgia Tech

What is Computation?

- Well-defined sequence of state changes (with purpose)
- Computation state: memory contents, input, output.

- Completeness: Church-Turing “thesis”: Anything Computable can be Computed by a Turing Machine.

- TM = finite-state machine + tape

Computation is Universal

- Planetary systems
- Weather
- Metabolic networks in your body
- Ant colonies
- Brains

The Brain, on a slide

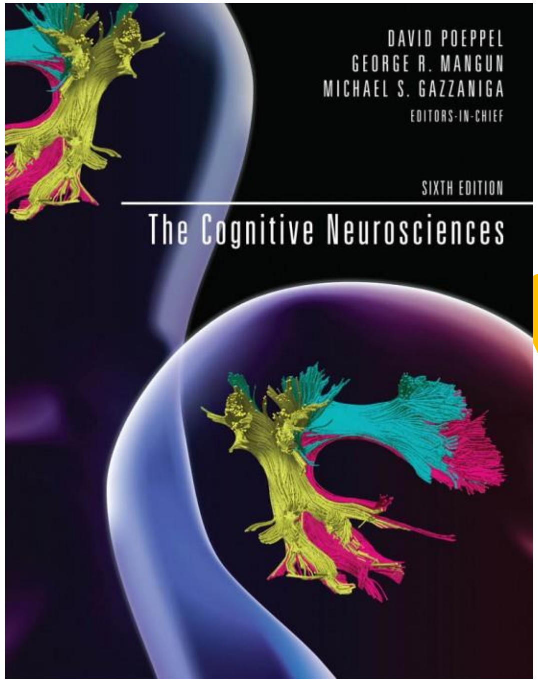
- A network of ~ 80 billion neurons
- $\sim 10^3 - 10^4$ connections per neuron
- Synapses (connections) have strengths, new synapses can form and existing ones might disappear during life
- Individual neurons "spike"/fire based on activation rules that are functions of signals on their input synapses.
  - Nonlinear activations; a common model is a threshold function
  - But there are >1000 types of neurons
  - Signals have a temporal aspect, with firing "rates" and "patterns".

How does the Mind (perception, cognition, …) emerge from the Brain (neurons, synapses, …)?

How does the Mind emerge from the Brain?

Despite accelerating progress in neuroscience
and increasing insight in cognitive science,
an overarching theory remains elusive

How does the Mind emerge from the Brain?



Mind the Gap

"…we do not have a logic for the transformation of neural activity into thought … I view discerning [this] logic as the most important future direction of neuroscience…"

Richard Axel, *Neuron*, Sep 2018

What kind of formal theory would qualify as Axel's "logic"?

- Hopfield Nets
  - A bit specific to storage and retrieval…
- Les Valiant's neuroidal model [1995]
  - allows for general operations on synapses (e.g., arbitrary state changes)

What are we looking for?

- a computational system (why?)
- consistent with current understanding of the brain
- explains cognitive phenomena
- ***what are the basic data types and the basic operations?***

Computation in the Brain: What is the right level?

- Whole brain?

?

- Spiking neurons and synapses?
- Dendrites?
- Molecules?

NEMO **(NEural MOdel)**

- A formal probabilistic model of the brain
- One basic data type
- A few elementary operations
- A completeness theorem
- A killer app ***(language!)***

Papers

Papadimitriou, Vempala, Mitropolsky, Collins, Maass (2020):
“Computation by assemblies of neurons,” PNAS.

Jung, Mitropolsky, Papadimitriou, Vempala (2021):
“An Interactive Tool for Computation with Assemblies of Neurons”, NeurIPS Demo

Mitropolsky, Collins, Papadimitriou (2021):
“A Biologically Plausible Parser”, TACL.

Dabagia, Papadimitriou, Vempala (2022):
“Assemblies of neurons can learn to classify well-separated distributions,” COLT.

Reid, Vempala (2023):
“The k-Cap Process on Geometric Random Graphs,” COLT.

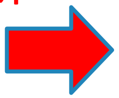
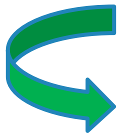
Dabagia, Papadimitriou, Vempala (2023):
“Computation with Sequences in the Brain,” arXiv preprint.

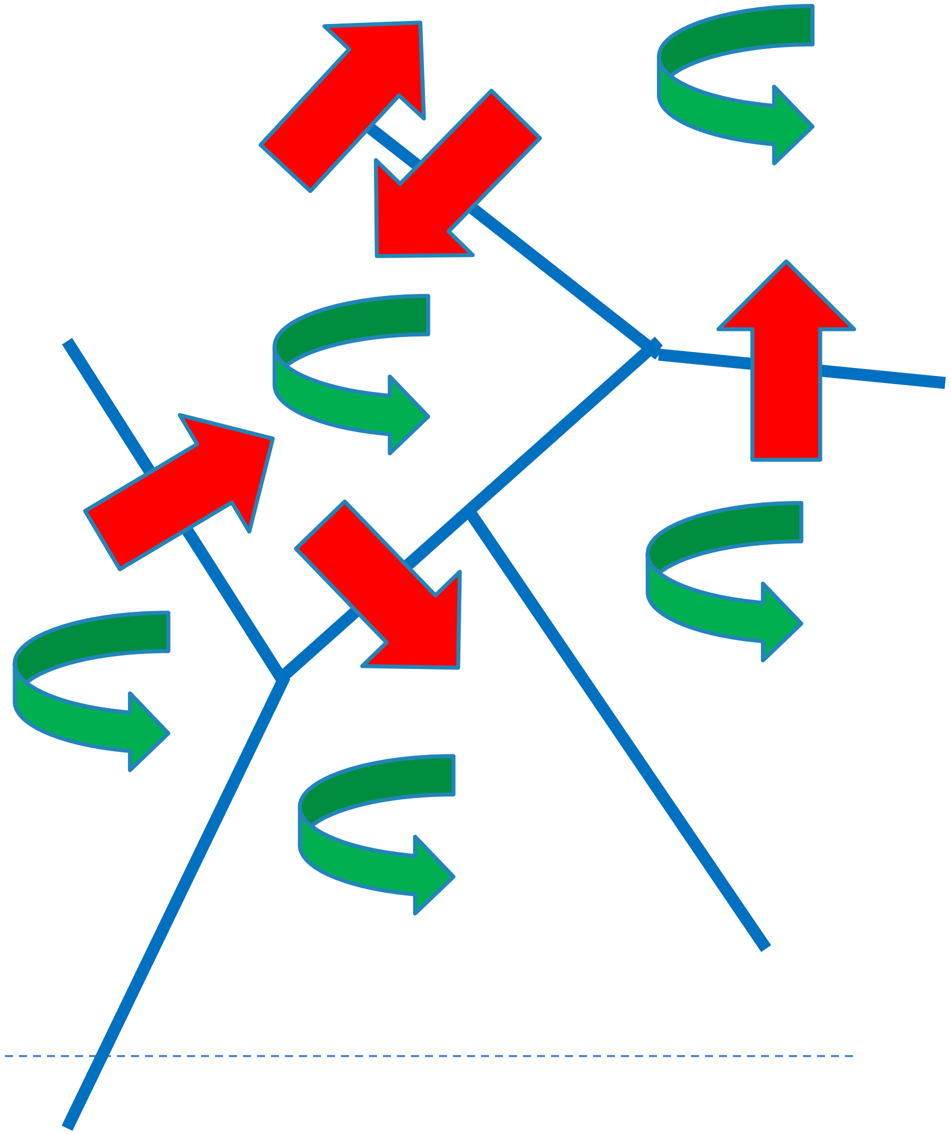
Random Graph

- $G_{n,p}$: A graph on n vertices, where each pair of vertices $\{i, j\}$ is independently assigned an edge with probability p.
  - $p = 0$: empty graph
  - $p = 1$: complete graph (clique)
- A great basic model with surprising structure.
  - Max degree is concentrated near its expectation, so is max clique.
  - As p increases from 0 to 1, any edge-monotone property goes from being unsatisfied to satisfied in some narrow interval of length $o(1)$, i.e., it has a "sharp" threshold.
  - E.g., connectivity, existence of a matching, Hamiltonian cycle etc.

- Directed $G_{n,p}$: each possible arc (i, j) is present independently with probability p.
- A simple, unrealistic but useful model of the connectome.

A Mathematical Model of the Brain

- Finite number of brain regions
- Each contains n neurons
- **Inhibition**: only k fire (selection)
- Some pairs of areas are connected by directed bipartite $G_{n,p}$ 
- All are recurrently connected by directed $G_{n,p}$ 

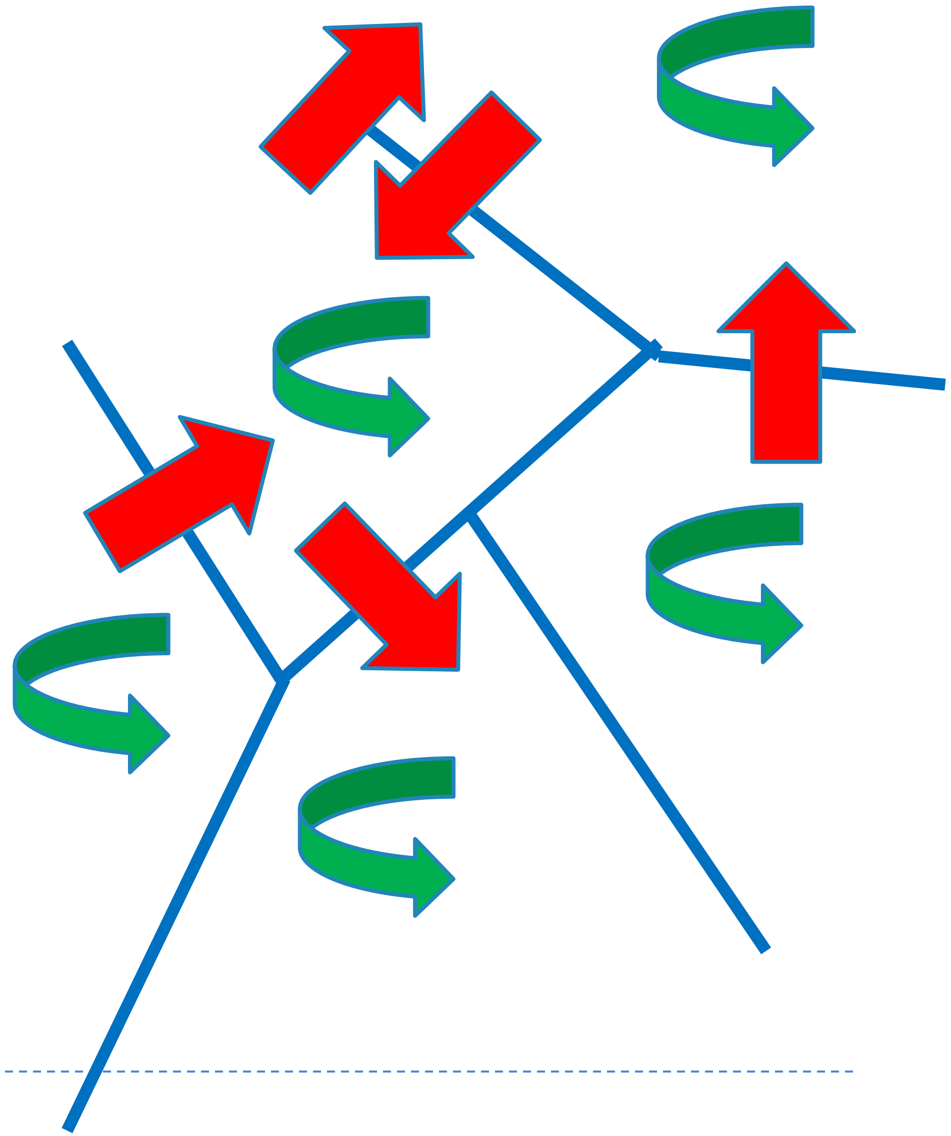


The Model (cont.)

- Neurons fire in ***discrete steps***
- The k neurons with highest input fire ***(RP&C)***
- Connections between areas can be inhibited/disinhibited

- **Plasticity** β: If i fires, and then j fires in the next step, the weight of synapse $i \rightarrow j$ is multiplied by $(1 + \beta)$

- Also: homeostasis,…



How realistic is it?

- Reasonably so, for a formal model of the brain
- The **discrete steps** assumption is certainly unrealistic, but likely not distortive
- Plasticity and assemblies used in a **rapid time scale**
- A productive compromise between realism and usefulness, allowing us to be rigorous

- Q: will computation and learning ***emerge***, rather than have to be “programmed”?

So, back to the question:

- What is the basic data type?
- (larger than neuron, smaller than brain)?

Assemblies (or *ensembles*) of neurons

- [Hebb 1949, Harris 2003, 2005; Buzsaki 2008, 2010, Yuste et al. 2019,…]
- Assembly: A large ***and densely intraconnected*** set of excitatory neurons in a brain area whose firing (in a pattern) is tantamount to our thinking of a particular memory, concept, person, name, word, episode, etc.
- **G. Buzsaki 2020, *The Brain Inside Out:***

  "assemblies are the alphabet of the brain"

The Assembly Hypothesis

- There is an intermediate level of brain computation
- Call it the Assembly Calculus
- It is implicated in higher cognitive functions such as reasoning, planning, language, story-telling, math, music, …
- Assemblies of neurons are its basic representation – its “data type”

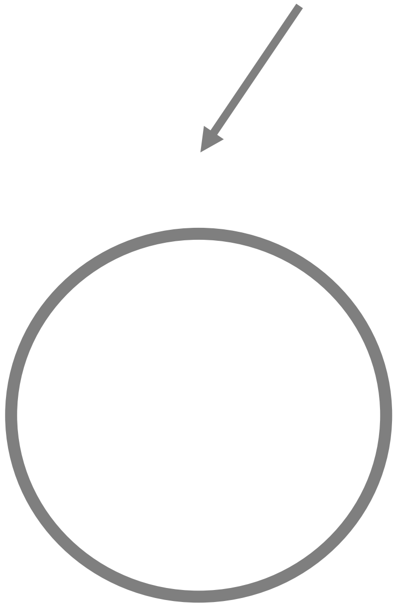
The Assembly Hypothesis (cont.)

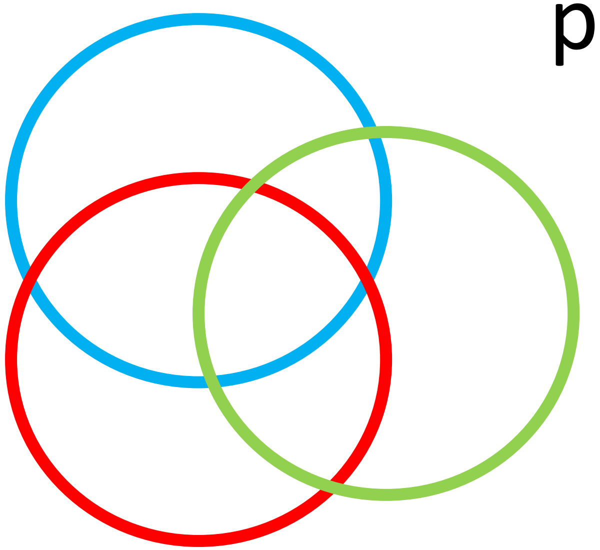
What are the operations of the Assembly Calculus?

- Project
- Associate
- Pattern complete
- Merge
- Bind (or reciprocal project)
- ...plus a few control commands

Projection of an assembly: **(RP&C)***

Assembly





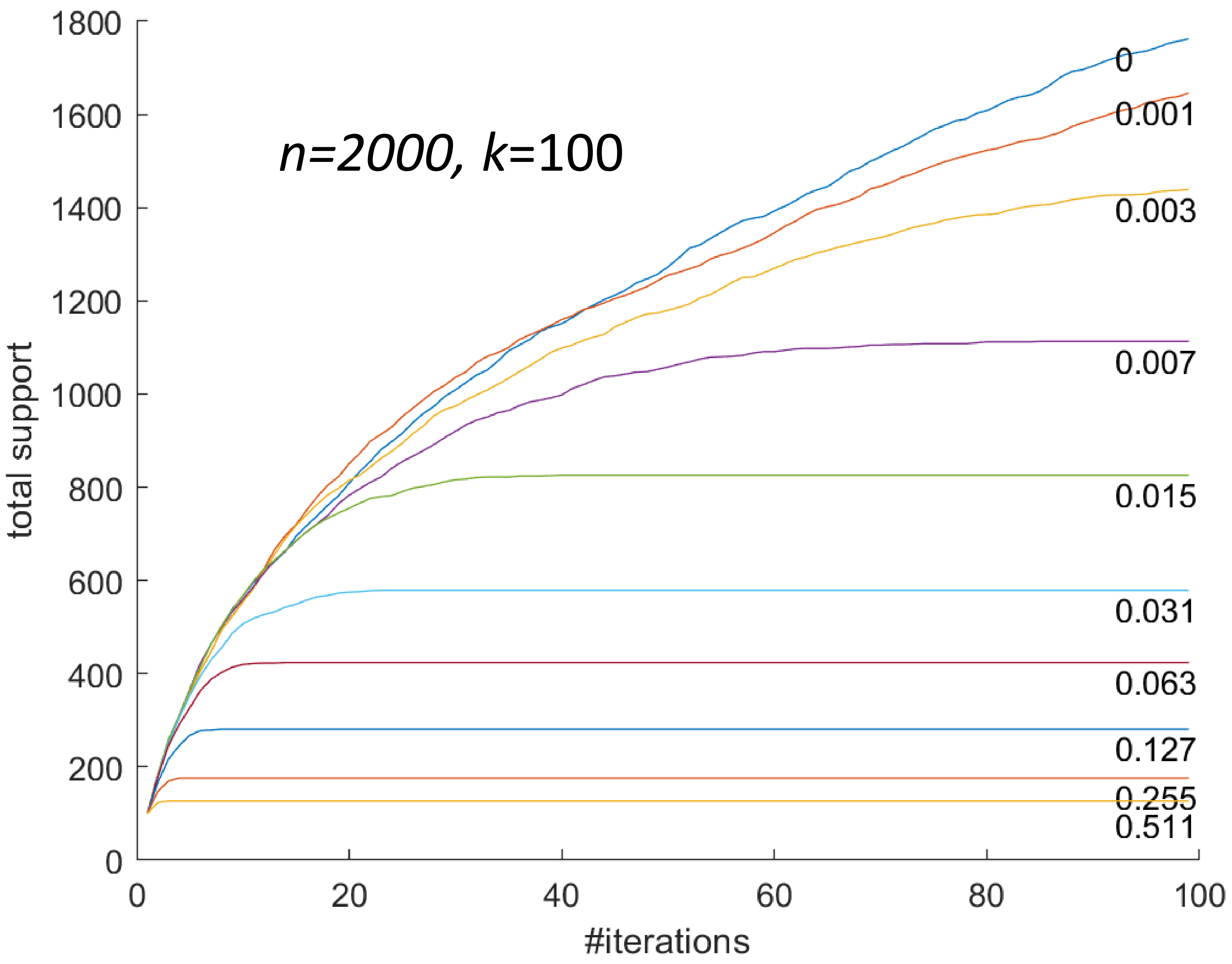
Also, Hebbian plasticity

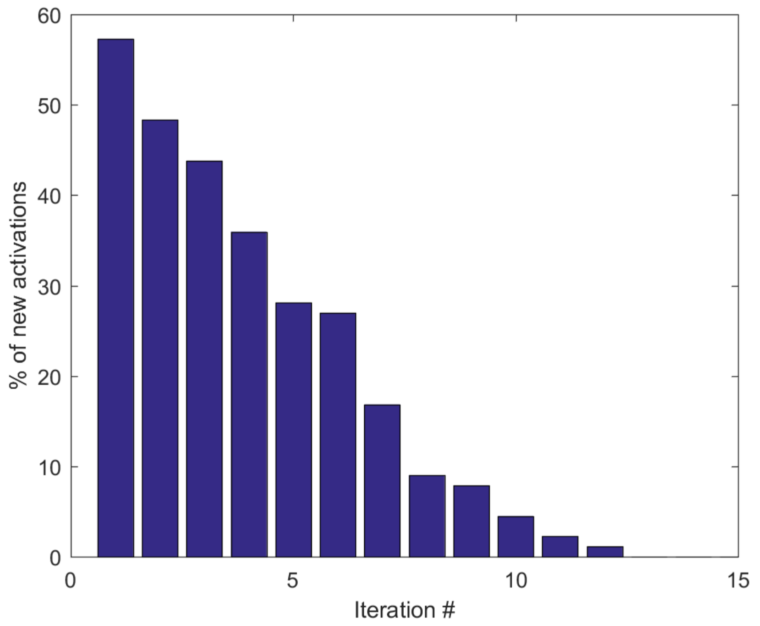
Convergence of **(RP&C)*** to an Assembly

Theorem (PV2018): The projection process converges exponentially fast, with high probability, and the *total number of cells involved* is at most:

- $k + o(k)$ if $\beta \geq \beta^*$
- $k \cdot \exp(0.17 \cdot \ln(n/k)/\beta)$ if $0 < \beta < \beta^*$:

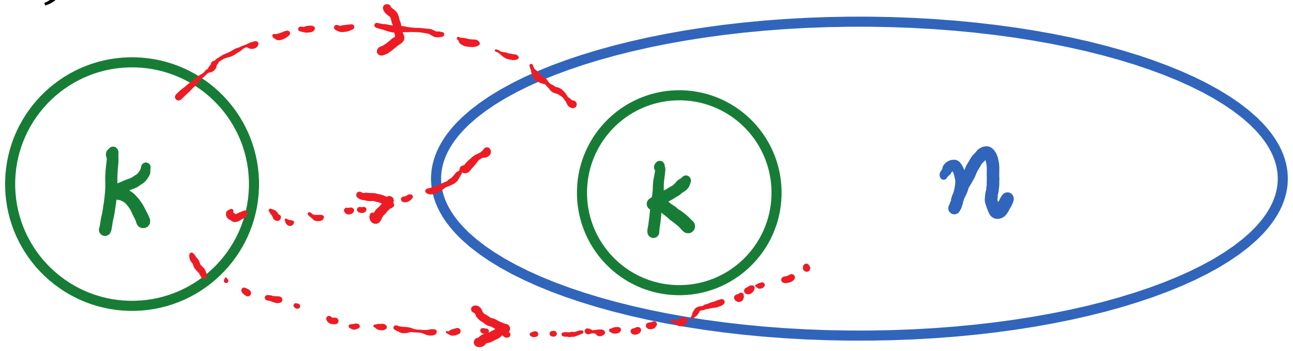
- $\beta^* = \dfrac{\left(\sqrt{2}-1\right)}{1+\sqrt{\dfrac{pk}{\ln(n/k)}}}$


n=2000, k=100
total support
#iterations
0
0.001
0.003
0.007
0.015
0.031
0.063
0.127
0.255
0.511
0
200
400
600
800
1000
1200
1400
1600
1800
20
40
60
80
100


% of new activations
Iteration #
0
10
20
30
40
50
60
5
15

Proof of Convergence

- First cap A_1: k neurons receiving highest input from stimulus, i.e., top k of n draws from $N(pk, pk)$



- Second cap: k neurons receiving highest input from **stimulus+A_1**, i.e., top k of n draws from $N(2pk, 2p(1-p)k)$
  - Competition between k previous winners and much larger pool.
  - Let $\mu_1, \ldots, \mu_t$ be the fraction of new winners at each step. Then,

$$C_1 = pk + \sqrt{2pk \ln \frac{n}{k}}, \qquad C_t = 2pk + 2\sqrt{pk \ln \frac{n}{\mu_t k}}, \qquad t \geq 2.$$

Proof of Convergence

$$C_1 = pk + \sqrt{2pk \ln \frac{n}{k}}, \qquad C_t = 2pk + 2\sqrt{pk \ln \frac{n}{\mu_t k}}, \qquad t \geq 2.$$

- For Step 2, for a neuron to stay in the cap, it suffices to have:

$$(1+\beta)C_1 + X \geq C_2 \quad \text{with} \quad X \sim N(pk, p(1-p)k)$$

where the $(1+\beta)$ is the plasticity boost for previous round winners

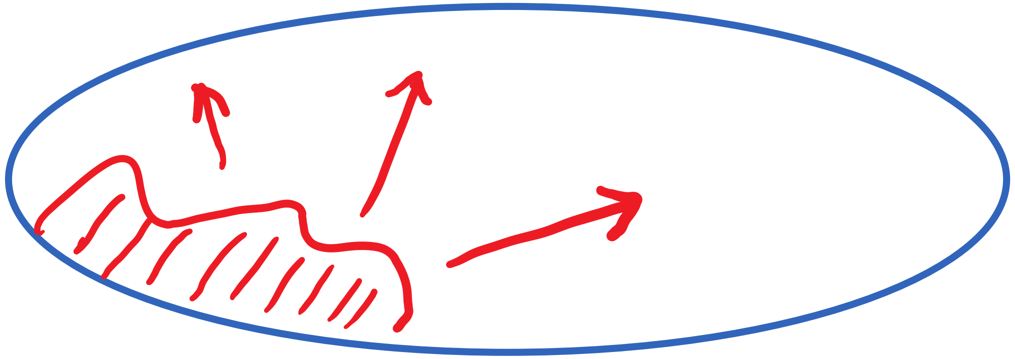
- This determines the threshold β^*
- Winners are more and more likely to survive: $(1+\beta)^i C_i + X \geq C_{i+1}$

Pattern Completion

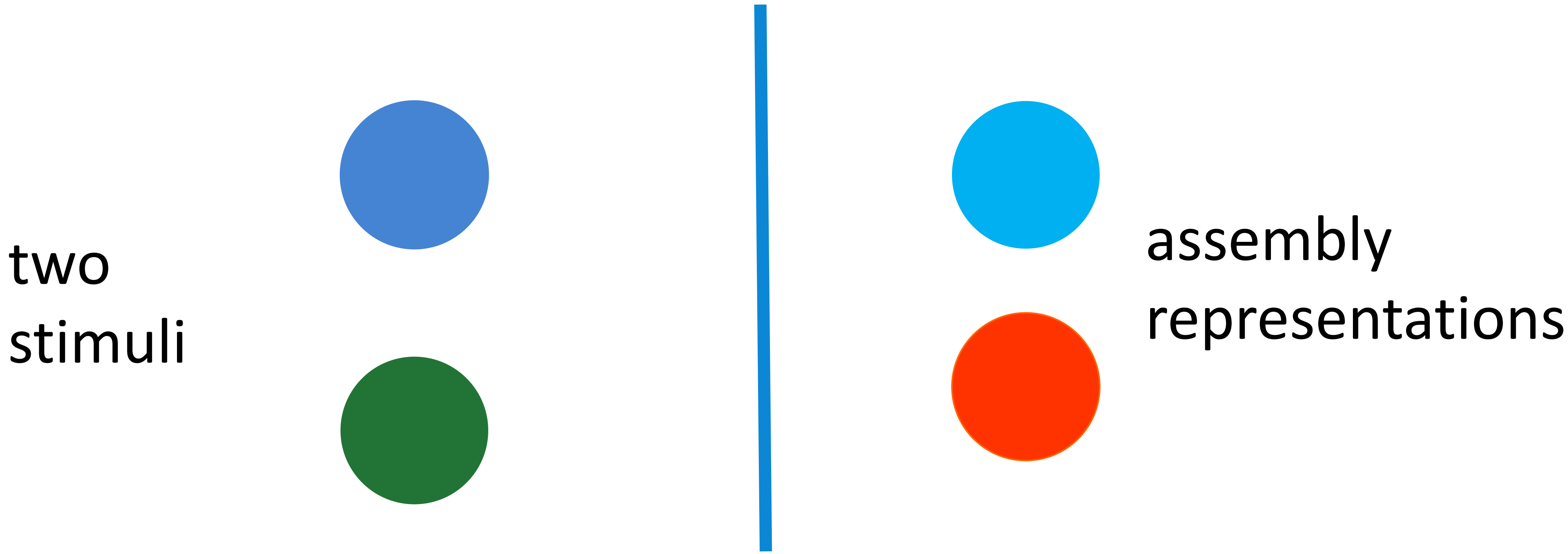
- Recall: Future presentation of the same or *similar* stimulus fires mostly the same assembly.
- What about firing a subset of the assembly itself?
- Yes! Suffices to fire fraction of created stimuli to complete to (almost) the rest.

- Theorem [Collins-Mitropolsky-P.-V. 2019]. For any $\varepsilon \in (0,1)$, with $T > T(\varepsilon)$ presentations of a stimulus, the resulting assembly A has the property that firing ε fraction of the assembly results in completing to 90% of A.

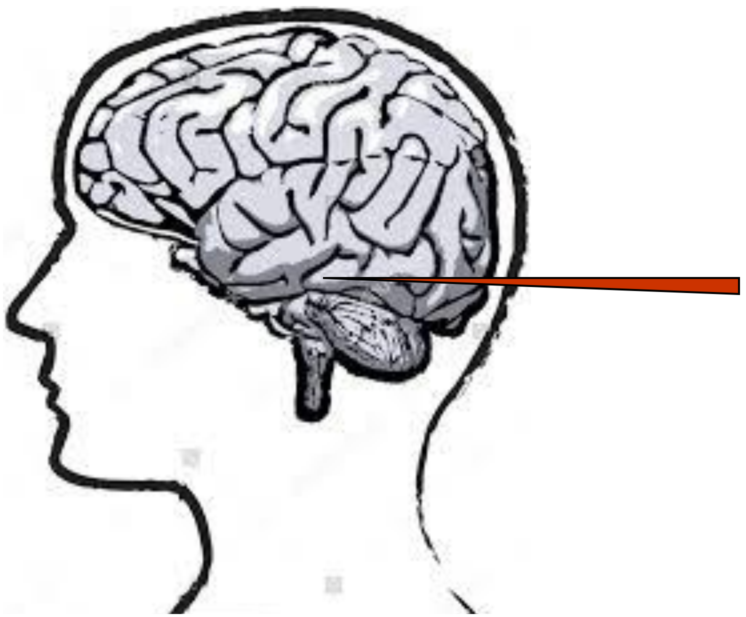
- Benefit of recurrent connections!



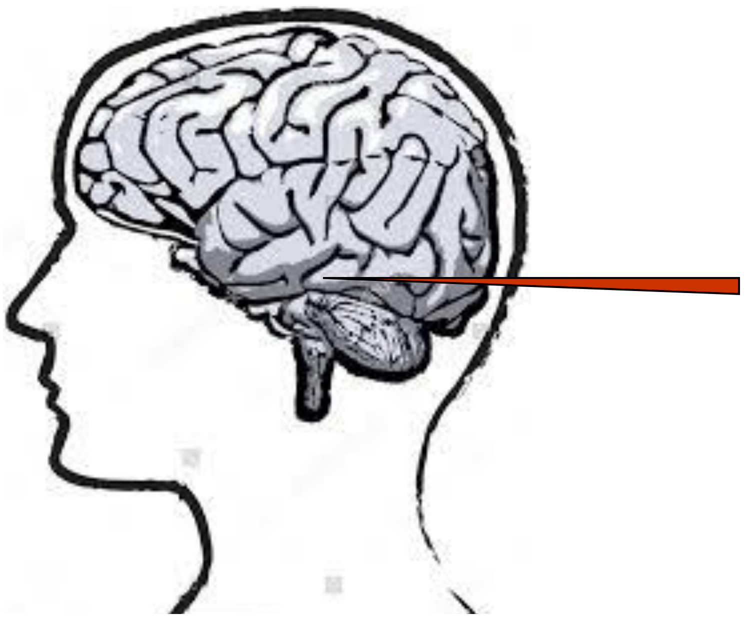
Association of Two Assemblies



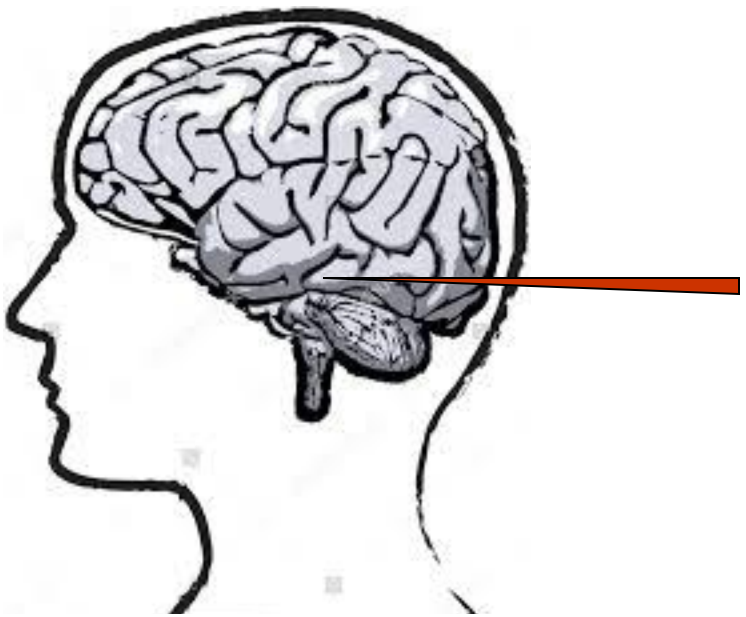
The [Ison et al. 2016] experiment

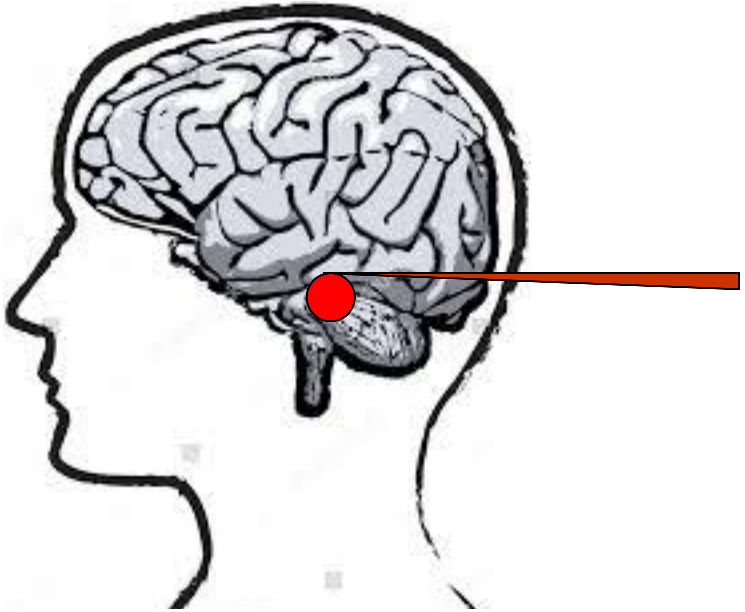


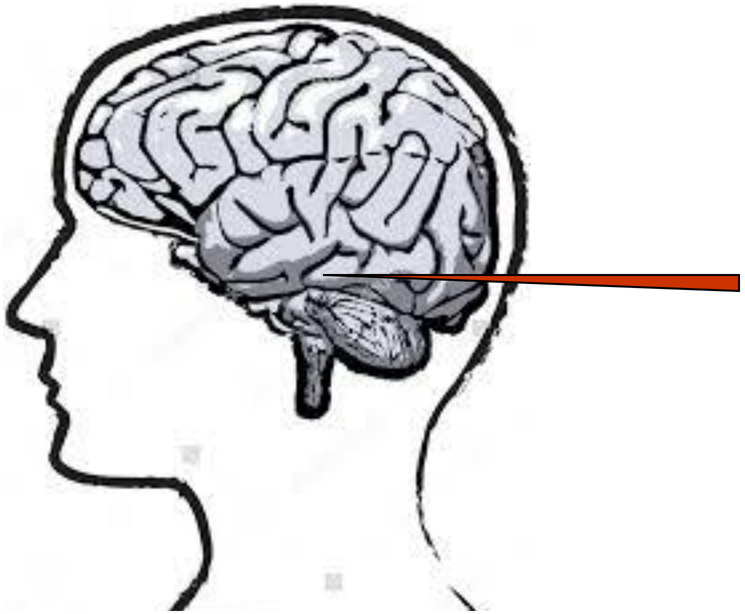
The [Ison et al. 2016] experiment

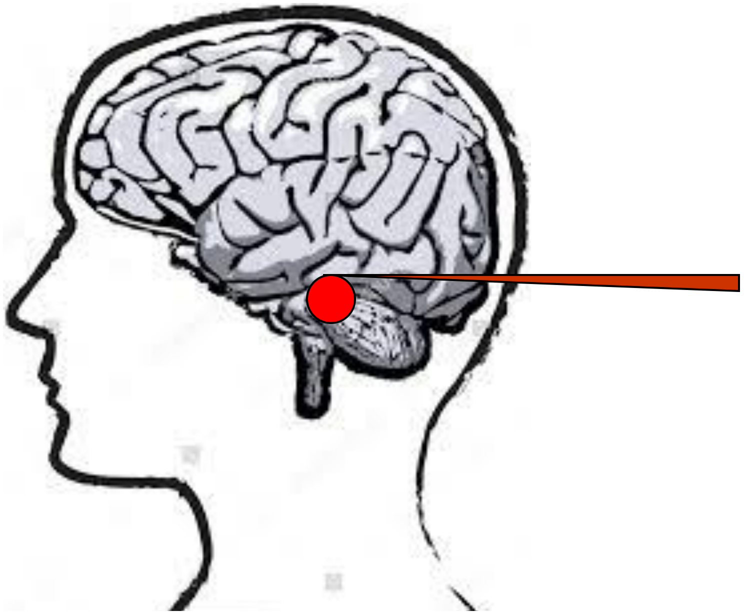


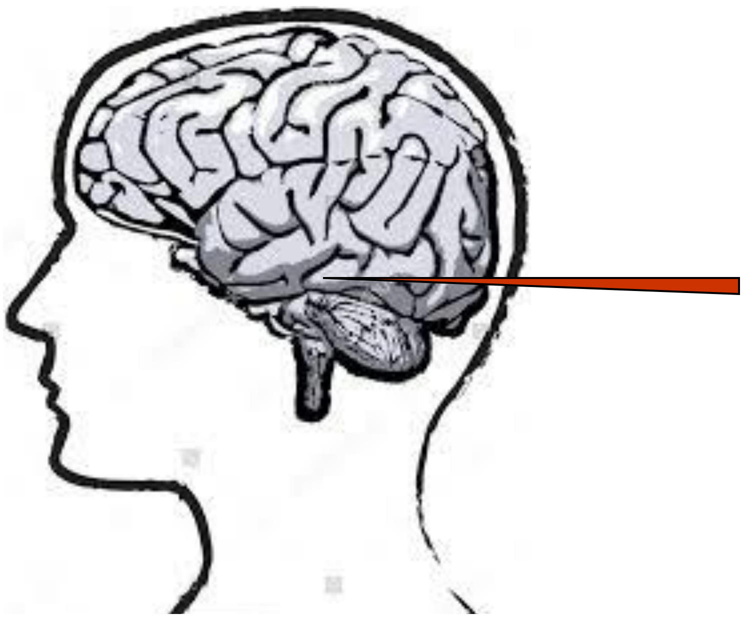
The [Ison et al. 2016] experiment

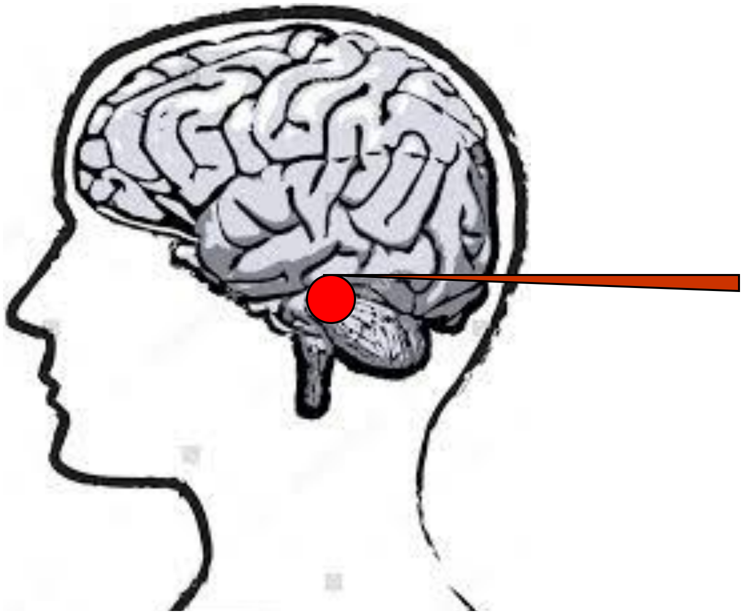


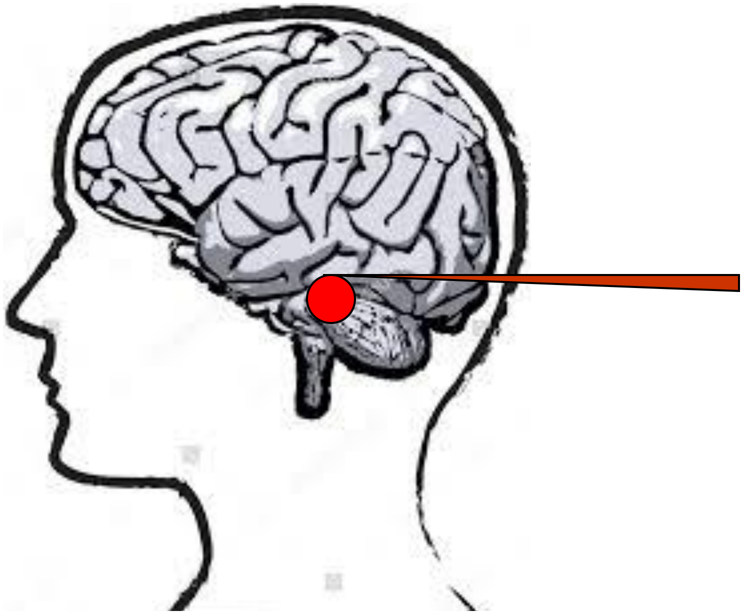


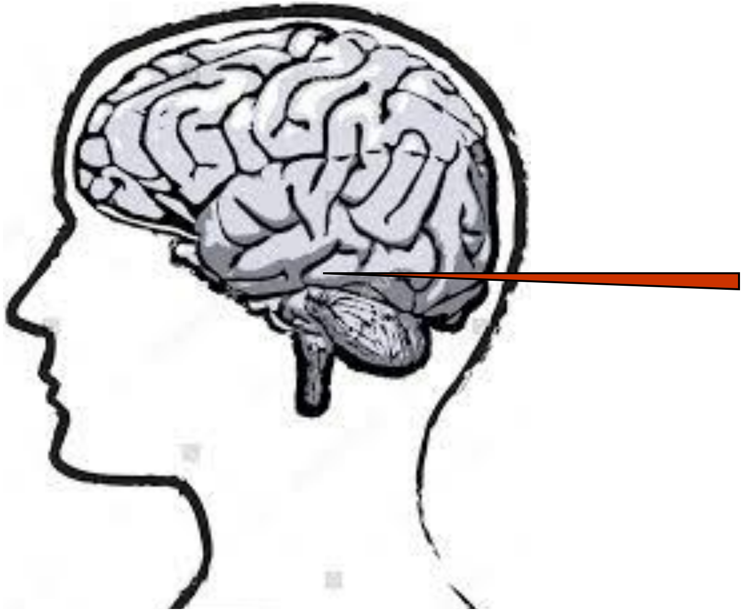


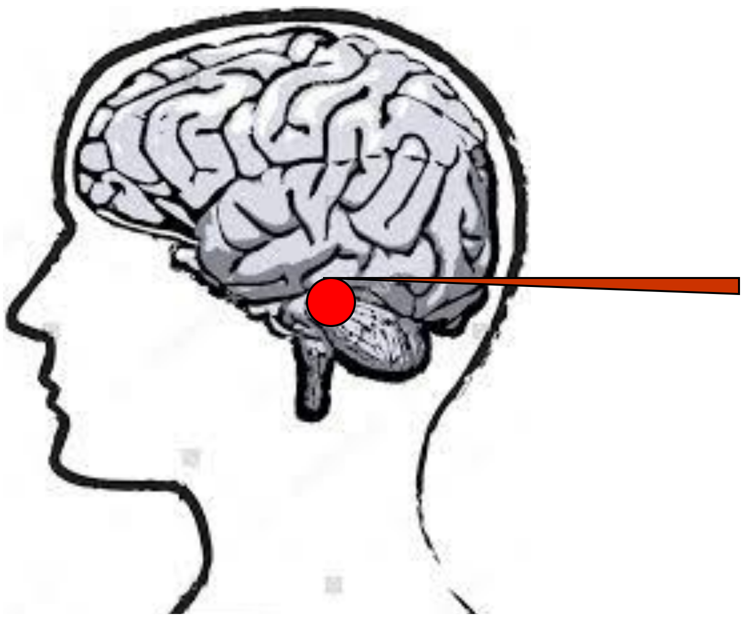








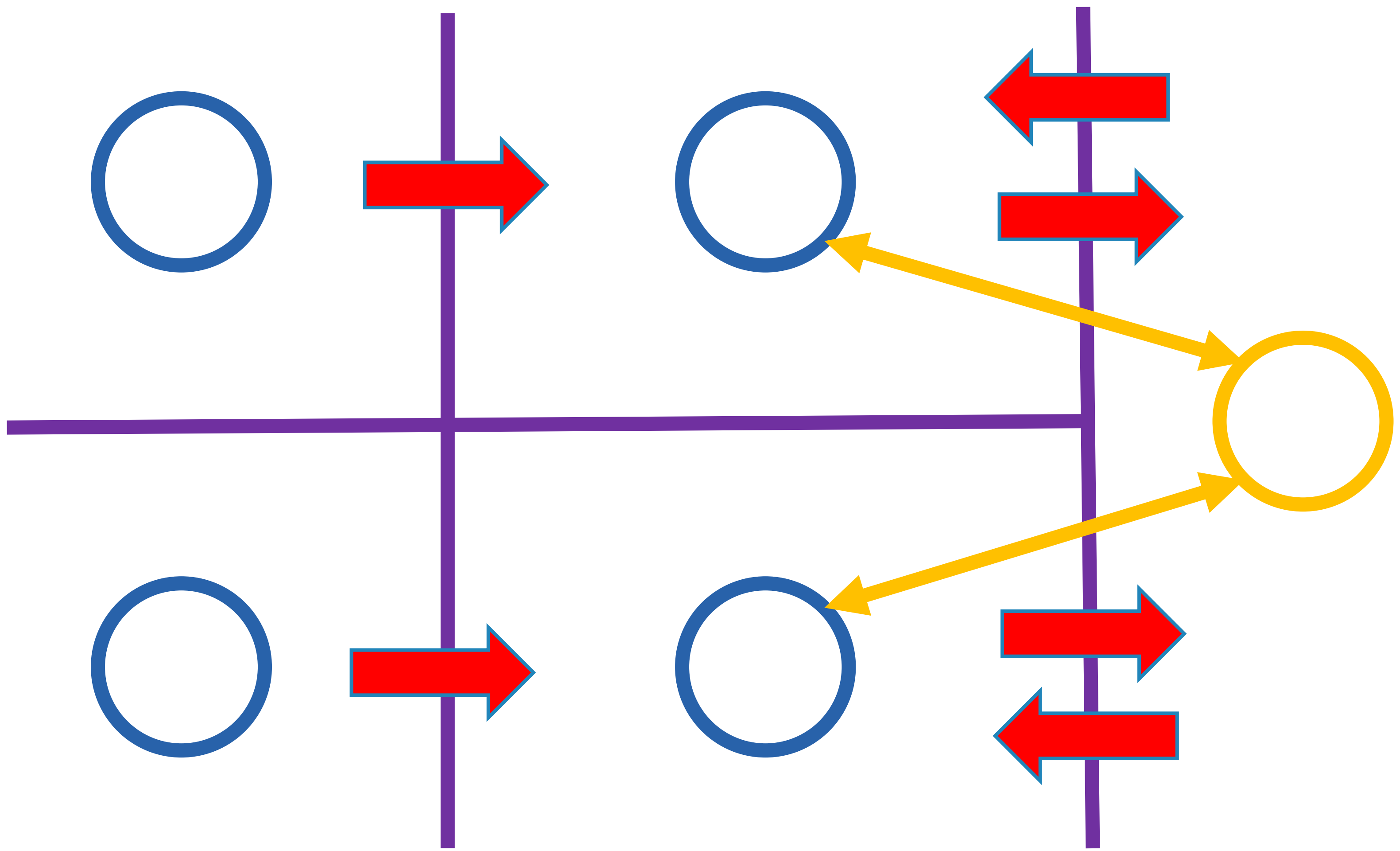




Association

- Provably happens with the $G_{n,p}$ assumption

Merge: a more complicated operation



…and useful for Language (Berwick-Chomsky)

Assembly Calculus recap

- `project`(y, B, x)
- `associate`(x, y)
- `pattern_complete`(x, y)
- `merge`(x, y, B, z)
- and control commands: `activate`(x), `read`(), `disinhibit`(A)

Q: ***How powerful is this system?***

Thm: ***It can perform arbitrary $\sqrt{n}$-space computations.***

In the rest of the talk

- Learning with assemblies
- Sequences!
- Giving up control
- Moving away from $G_{n,p}$
- Language!

Learning: How do brains learn?

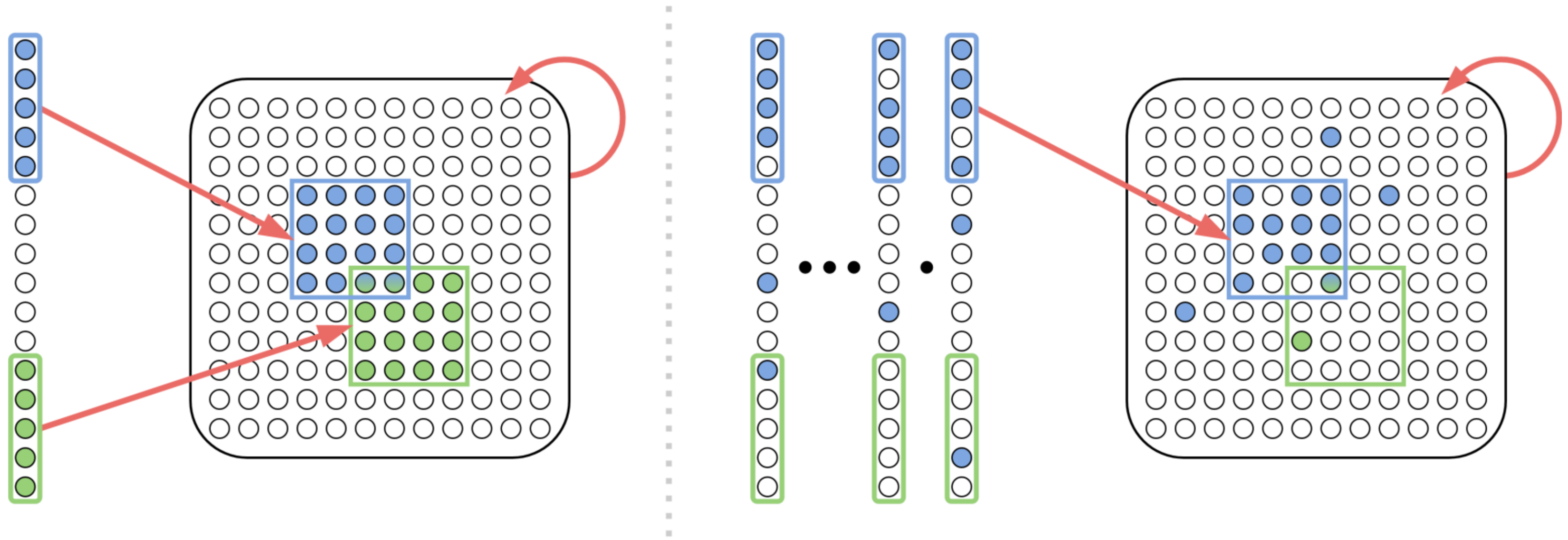
- **Crucial problem** for NEMO / Assembly Calculus

- Locally: plasticity
- Globally: does the brain do Gradient Descent (GD)?
  - A topic of debate, with little evidence that it does.

- Is synaptic plasticity an effective learning mechanism?

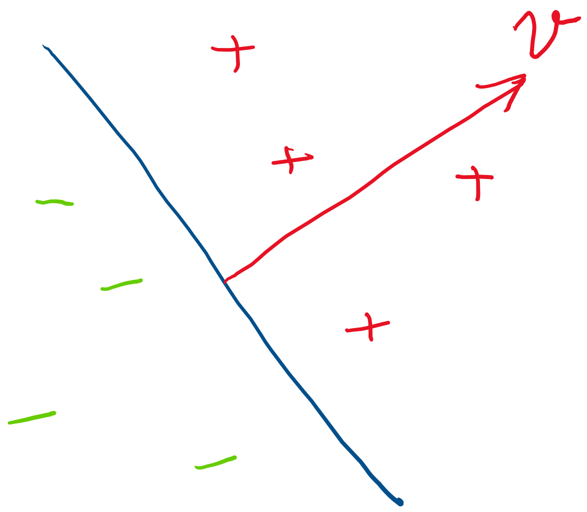
Unsupervised Learning with Assemblies

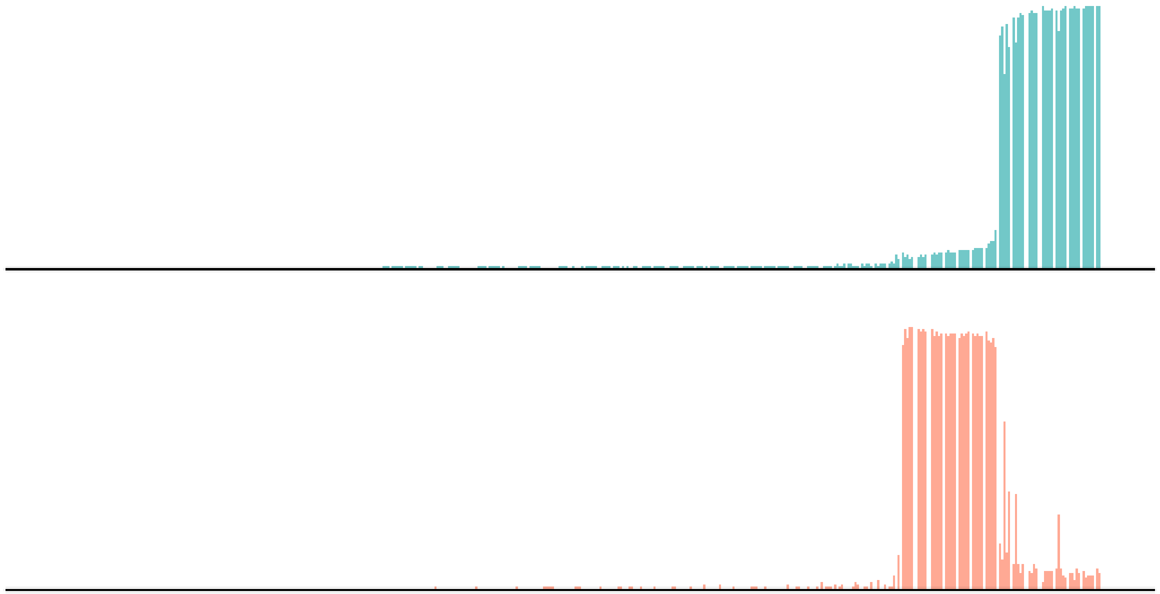
- **Question:** We know that the assembly projection process preserves overlap. But can it create assemblies to represent clusters of data, i.e., higher-level assemblies?

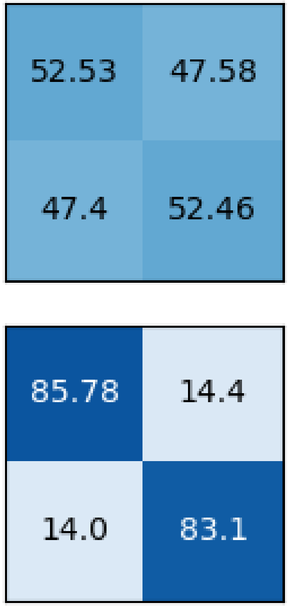


Mildly supervised Learning of Halfspaces?

- Examples from halfspace and its complement: x Bernoulli, $x \cdot v \geq E(x \cdot v) + \Delta$
- 5 examples from each, presented consecutively







- Overlap between projections much higher than between corresponding assembly representations.
- Question: Can such halfspaces be provably learned by assemblies? Yes!

Internally Generated Cell
Sequences in the Rat Hi

Eva Pastalkova, Vladimir Itskov,* Asohan Amarasingham, Gy

A long-standing conjecture in neuroscience is that aspects of cog
to self-generate sequential neuronal activity. We found that relia
assemblies in the rat hippocampus appeared not only during sp
absence of changing environmental or body-derived inputs. Dur
task, each moment in time was characterized by the activity of a
Identical initial conditions triggered a similar assembly sequenc
gave rise to different sequences, thereby predicting behavioral
sequences were not formed in control (nonmemory) tasks. We h

LETTER

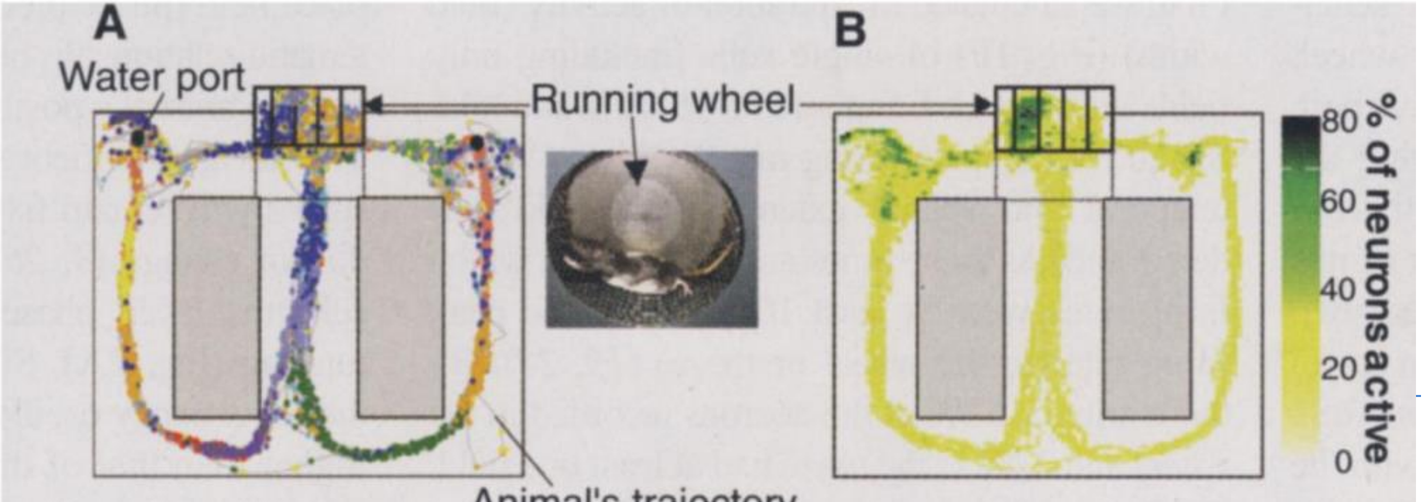
doi:10.1038/nature09633

Preplay of future place cell sequences by hippocampal cellular assemblies

George Dragoi[1] & Susumu Tonegawa[1]

During spatial exploration, hippocampal neurons show a sequential firing pattern in which individual neurons fire specifically at particular locations along the animal's trajectory (place cells[1,2]). According to the dominant model of hippocampal cell assembly activity, place cell firing order is established for the first time during exploration, to encode the spatial experience, and is subsequently replayed during rest[3–6] or slow-wave sleep[7–10] for consolidation of the encoded experience[11,12]. Here we report that temporal sequences of firing of place cells expressed during a novel spatial experience occurred on a significant number of occasions during the resting or sleeping period preceding the experience. This phenomenon, which is called preplay, occurred in disjunction with sequences of replay of a familiar experience. These results suggest that internal neuronal dynamics during resting or sleep organize hippocampal cellular assemblies[13–15] into temporal sequences that contribute to the encoding of a related novel experience occurring in the future.

of template that were emitted during Fam-Rest were sorted by time, and spiking events were determined as explained above (subpanels b in Fig. 1). For each spiking event, we calculated a rank-order correlation between the novel arm templates and the temporal sequence of firing of the corresponding cells in the spiking events during Fam-Rest. The event correlation was considered significant if it exceeded the 97.5th percentile of a distribution of correlations resulting from randomly shuffling the order of place cells in the novel arm templates 200 times ($P < 0.025$). Forward[4] and reverse[3,4] preplay refers to the cases in which the sequence of place cells during Contig-Run and the firing order of the corresponding cells in Fam-Rest were in the same and opposite directions, respectively. In 91% of the preplay cases, the spiking events were correlated with the novel arm template in one direction only. The distribution of event correlation values obtained using the original novel arm templates was significantly shifted towards higher positive or negative values in comparison with the distribution of correlation



Sequences!

- Memorizing and Learning sequences of stimuli seems very natural to brains.

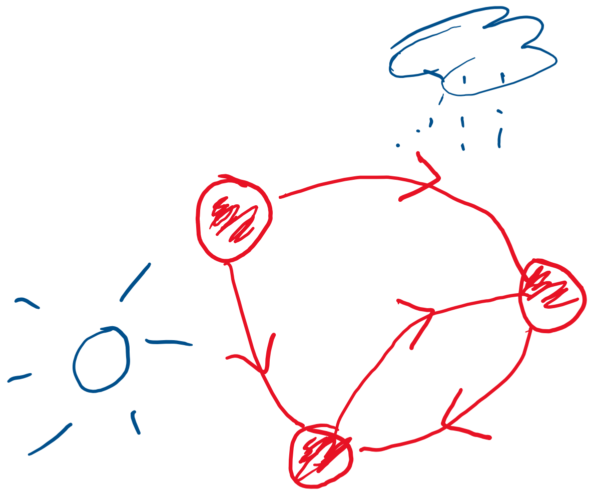
- Simplest sequence problem in the Assembly Model:
  - Present sequence of stimuli S1, S2, …, St in the input area (or activate assemblies A1, A2, …, At in area A) to area B.
  - Result should be projection of sequence to assemblies B1, B2, …, Bt in area B.
  - Activating any Si leads to activation of Bi followed by B(i+1),…. Bt.

- We'll try it soon!
- Thm [D-P-V'22]. This happens WHP in NEMO, after presenting the sequence $O(\log n)$ times.

Sequences!

Powerful consequences:

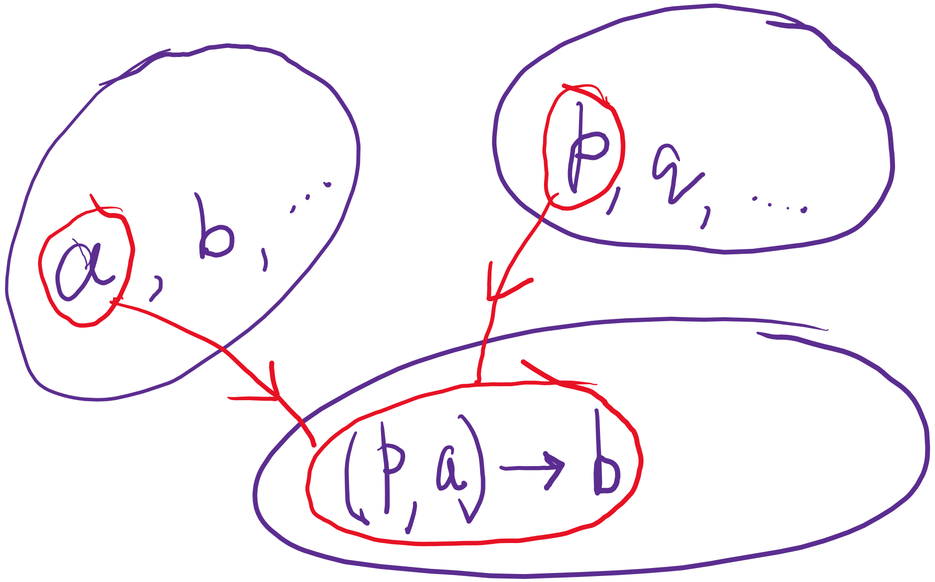
- Finite State Automata

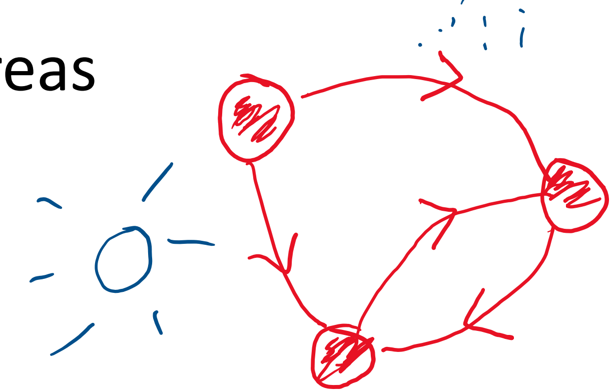


- Algorithms are finite-state machines; we learn them by simply being exposed to their rules.
- General-purpose simulation of Turing machine by simply presenting appropriate input sequence!

Brain computation without control commands

- Both FSMs and TMs operate in the AC *without control commands*
- Long-Range Interneurons whose firing can inhibit/disinhibit an entire area
- For FSM’s, we we just need alternation between firing areas

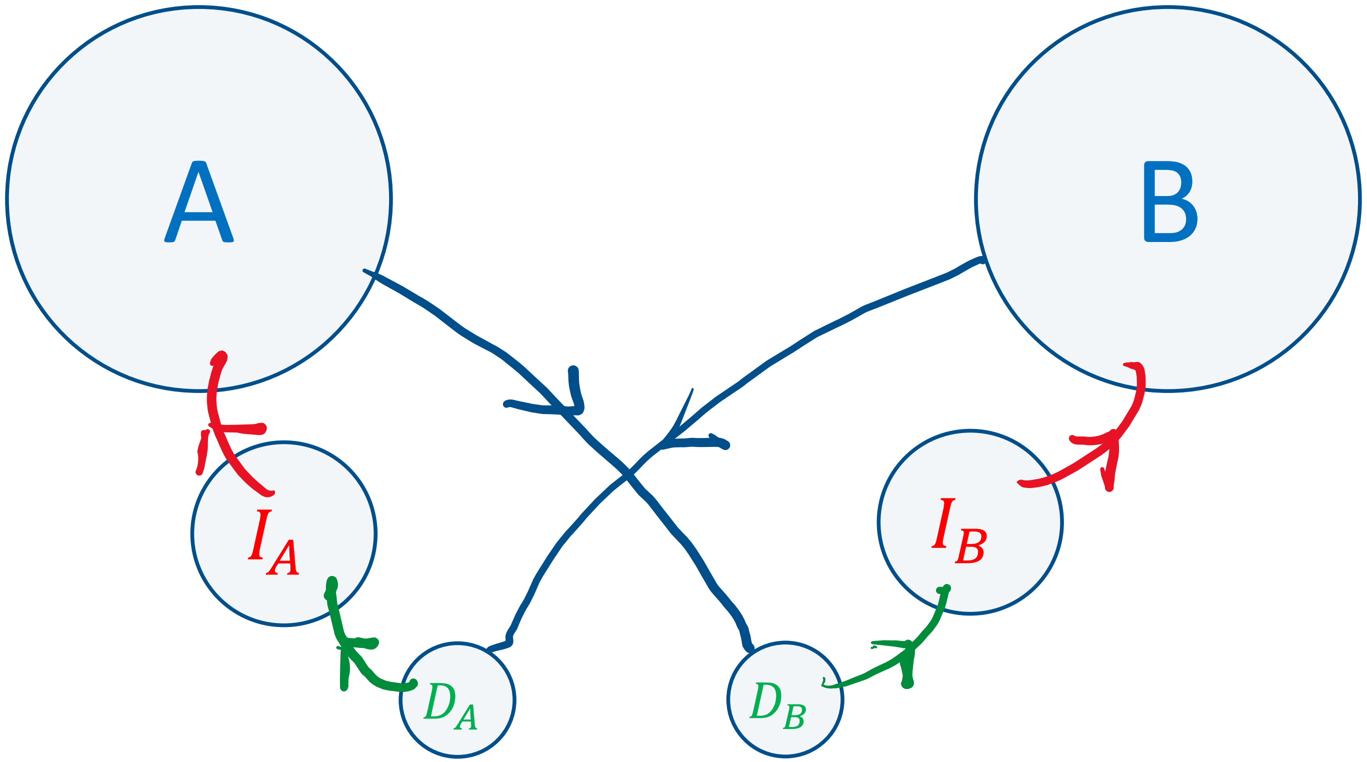




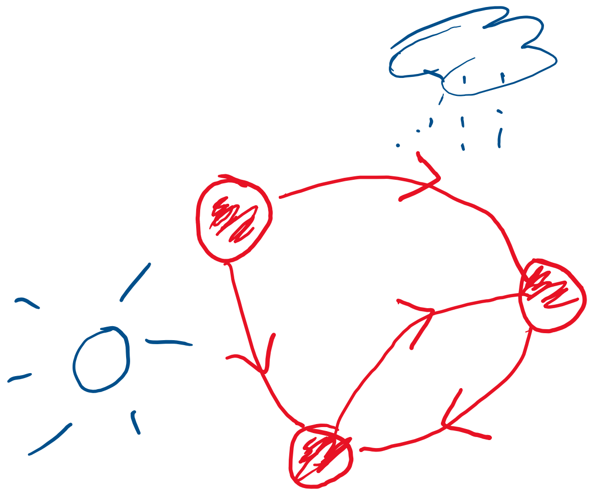
- General-purpose simulation of Turing machine ***without control commands*** by simply presenting appropriate input sequence!

The benefits of (long-range) inhibition

- Goal: make areas A and B fire alternately
- Populations of inhibitory neurons to inhibit: I_A, I_B and to *dis*inhibit: D_A, D_B



Brain computation without control commands



- The realization and execution of finite-state machines is **emergent** computation, *provably* --- an appropriate sequence of inputs, components of the AC operating by their local rules, and no overall control!

- https://github.com/mdabagia/learning-with-assemblies

Finally, Language!

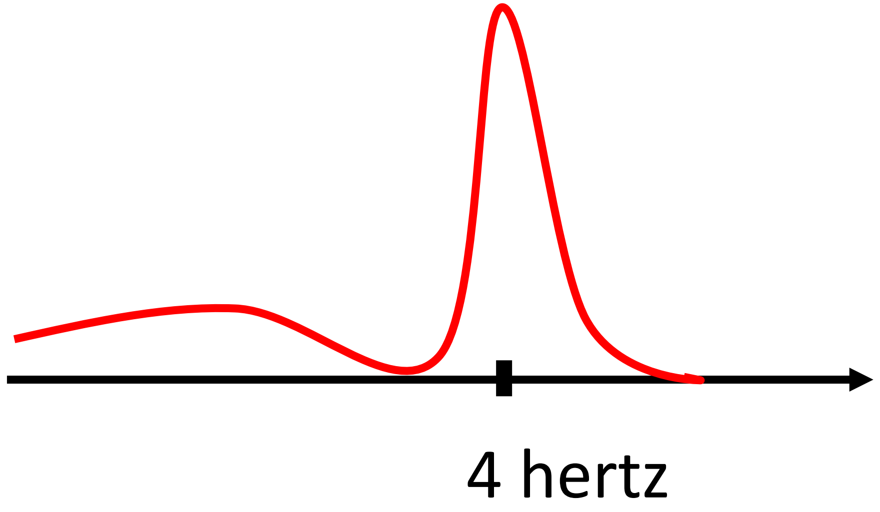
- A last-minute adaptation
- It evolved to exploit the brain's strengths
- Invaluable lens for studying the brain
- ***A deluge of recent experiments!***

…this text pulsates at four
Hertz, the rhythm of four
bits per second, and you
may find this rhythm a bit
familiar, because it coincides
with the rhythm of speech,
and I don’t mean my speech
or your speech, but speech
in general, by all speakers,
in all languages, worldwide.

…in contrast, this text pulsates a
dozen times faster at fifty Hertz,
the rhythm of spiking neurons
in the Brain…

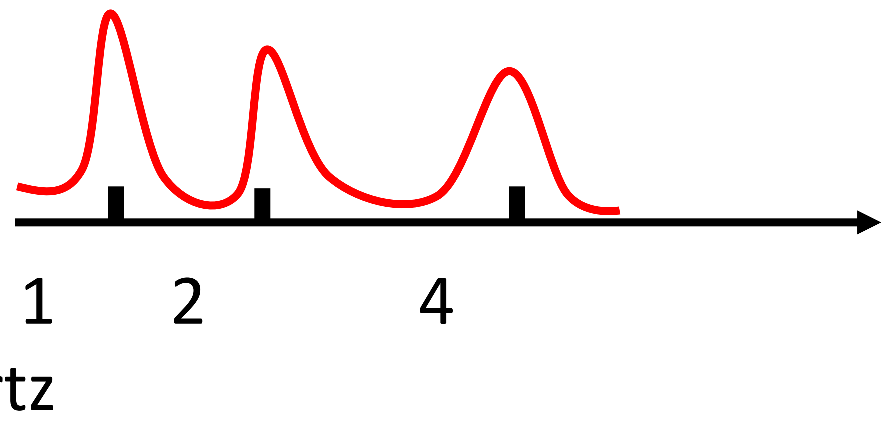
The [Poeppel 2016] experiment

fret ship hill give true melt fans blue guess hits then cats

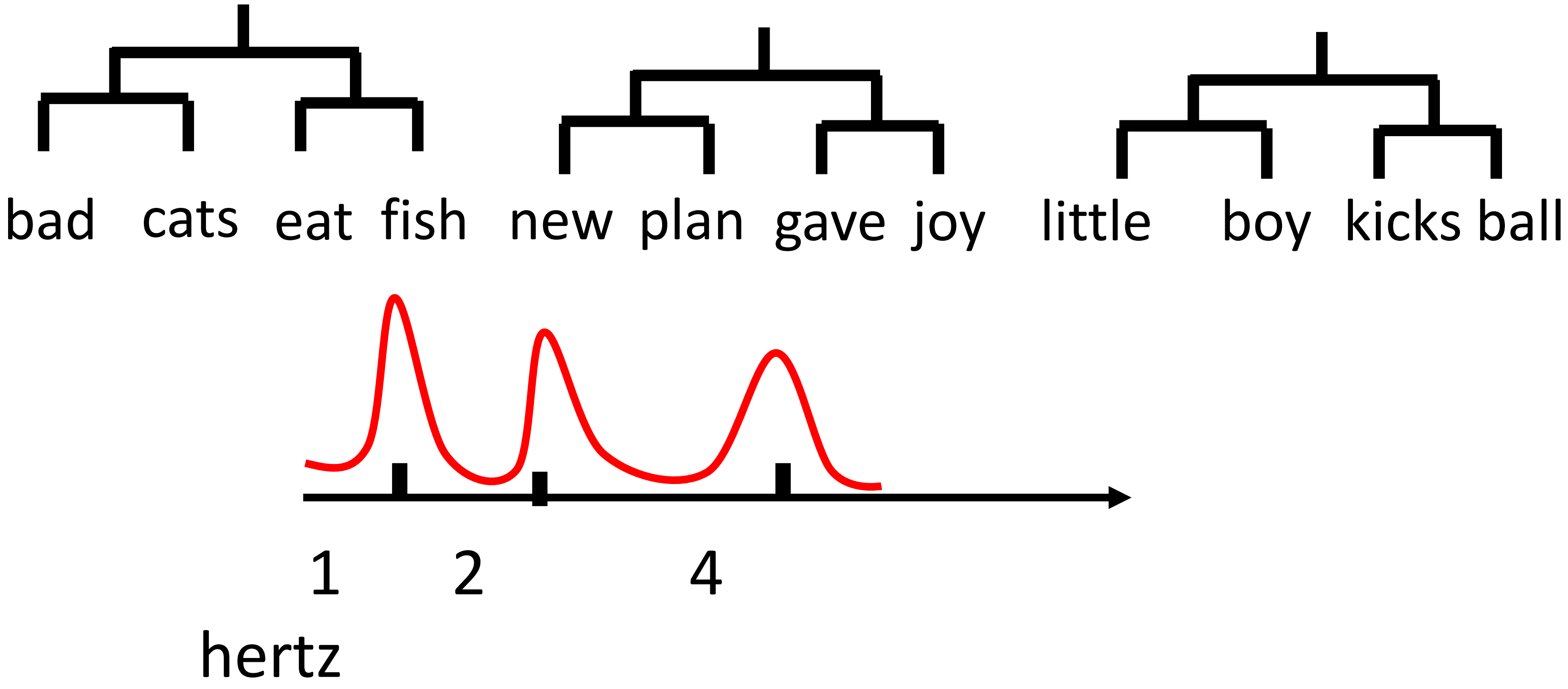


The [Poeppel 2016] experiment, stage II

bad cats eat fish new plan gave joy little boy kicks ball



Our interpretation



Zaccarella & Friedericci “Merge in the human Brain”

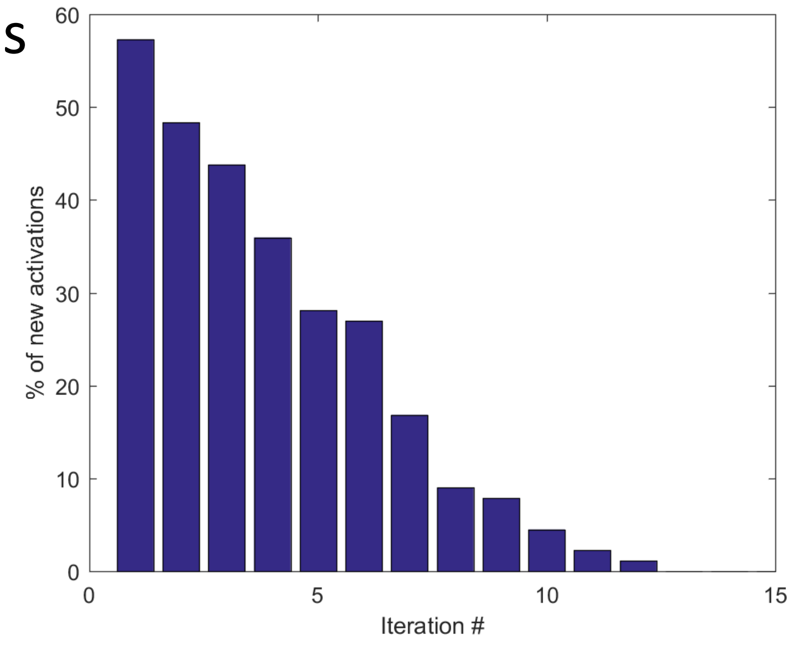
Front. Psych. 2015

- The completion of phrases, and especially of sentences, activates parts of Broca’s area

Assemblies and Language: Syntax Trees

Can each tree-building step be accomplished by about a dozen spikes?

Analytical results and simulations show that projection creates a new assembly by spiking repeatedly a few times! (~12)



Many Research Directions!

- On the Assembly Calculus
- Emergent computation: input and mechanism, no algorithm!

(removes one level of David Marr’s 3-level framework for brain)

- Random graph models of the connectome
- Sparsification, plasticity, and Brain Organization
- Learning with Assemblies
- Language: word embeddings, analogies and reasoning with neural assemblies, next-token prediction with NEMO

Open problems related to ***Assemblies and Convergence***

- **Assembly support size.** Is there a phase transition in the support size of an assembly from $\omega(k)$ to $k + o(k)$ as the plasticity parameter β increases?

- **What is an assembly?** We know the limit of a projection can oscillate between two overlapping sets. What are other possible (nonvanishing probability) limiting behaviors? Assemblies may have to be defined as distributions over [n] with k+o(k) support and firing patterns. How?

Nonrandom features of the Connectome

- [Song et al 2005]: reciprocity and triangle completion

of the data from 14 to 16-d-old animals when the majority of measurements were performed (see Figure S5). We found that bidirectional connections are also overrepresented in this subset of data. Results of other analyses that will be described later in the paper are also confirmed on this subset (Figure S5).

Finally, it is possible that some extreme degree of inhomogeneity in connections probability is able to explain the observed overrepresentation of reciprocal pairs, but this would probably reflect large local inhomogeneity in cortical connectivity patterns—possibly differences between subclasses [6,35], rather than experimental artifacts—and is in line with the main conclusions of this paper.

Three-Neuron Patterns

We extended our analysis by comparing the statistics of three-neuron patterns to those expected by chance [26,27]. We classify all triplets into 16 classes and count the number of triplets in each class. In order to avoid reporting over-represented three-neuron patterns just because they contain popular two-neuron patterns, we have revised the null hypothesis[26,27]. The control distribution was obtained numerically by constructing random triplets where constituent pairs are chosen independently, but with the same probability of bidirectional and unidirectional connections as in data (Figure 4A). For example, the actual probability of a unidirectional connection is (according to Figure 2B) 495/(3312 + 495 + 218) = 0.123. Then the probability of unidirectional connection from A to B is 0.123/2 = 0.0615, the same as from B to A (see Figure 4A). The probability of bidirectional connection is (according to Figure 2B) 218/(3312 + 495 + 218) = 0.0542. The probability of finding the particular triplet class in Figure 4A by chance is the product of the probabilities of finding the three constituent pairs and a factor to account for permutations of the three neurons.

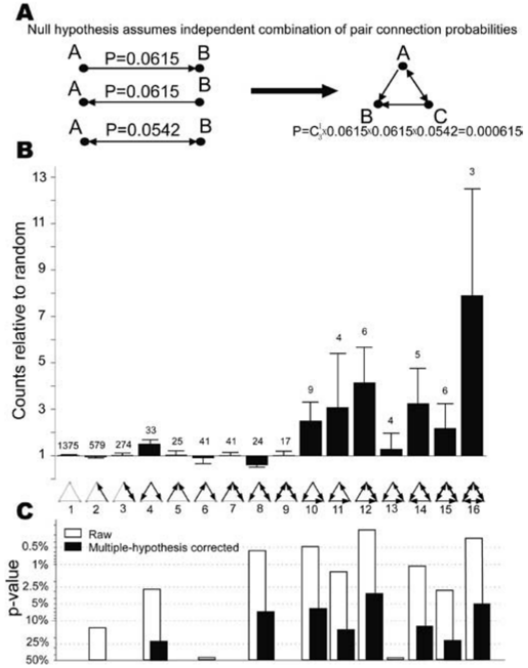


Figure 4. Several Three-Neuron Patterns Are Overrepresented as Compared to the Random Network

(A) Null hypothesis for three-neuron patterns assumes independent combinations of connection probabilities of two kinds of two-neuron patterns.

(B) Ratio of actual counts (numbers above bars) to that predicted by the null hypothesis. Error bars are standard deviations estimated by

Random Graph models

- $G_{n,p}$ with sufficient plasticity and edge density leads to stable assemblies and supports assembly operations
- The connectome shows noticeable deviations from $G_{n,p}$, e.g., in the number of triangles and other small subgraphs
- What deviations would be helpful? How to model them?

Departures from $G_{n,p}$

- They seem to **strengthen** the convergence results for Projection
- Reciprocity increases **assembly density**; triangle completion begets a **birthday paradox** phenomenon
- **Conjecture:** Assembly operations become provably more robust in the presence of these “departures.”

A Fundamental Convergence Problem

- Begin with k random vertices firing
- Repeat:
  - Vertex weight: sum of weights of edges coming from firing vertices
  - k vertices with highest weight fire.
- **Q: Does this process converge?**
- $G_{n,p}$ YES, but we need plasticity of edge weights.
- **Q:** on which graphs the process converges **without plasticity** (β = 0)?
- **A natural candidate:** Geometric random graphs

Geometric Random Graphs

- Vertices are assigned random values in an interval or square (could be a d-dimensional cube).
- Edge probability depends on distance between endpoints, e.g.,

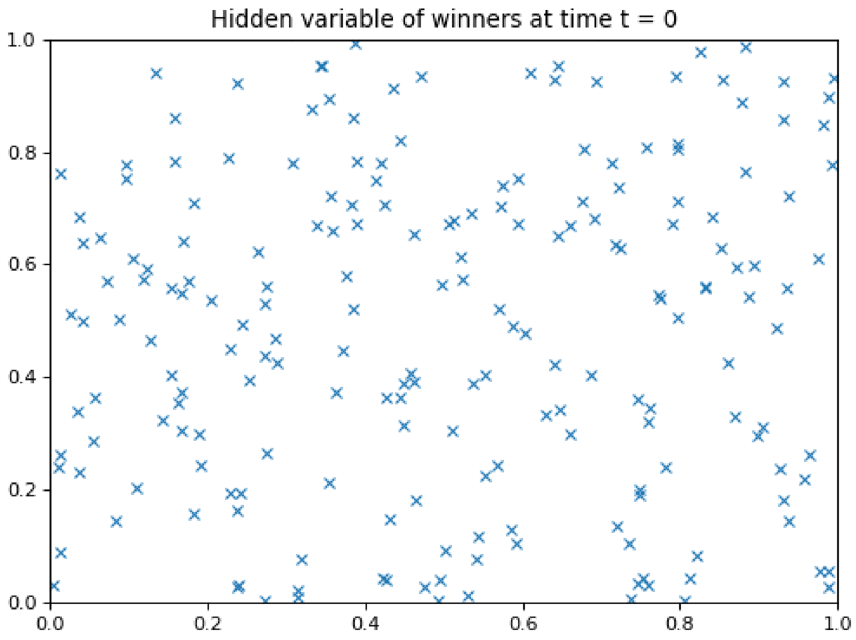
$$g(x, y) = \exp\left(-\frac{\alpha\|x-y\|^2}{2}\right)$$

Q. Does this model have an advantage over $\boldsymbol{G_{n,p}}$?

- Empirically, YES: Assembly projection appears to be more efficient with geometric random graphs and needs lower plasticity!

Convergence on Geometric Random Graphs

- Vertices are random points in a square, $n = 10{,}000, k = 200$.
- Edge probability is proportional to $e^{-\|x-y\|^2/2\sigma^2}$, edge density = 0.01.



- What are the fixed points?! And will there be convergence?

Convergence on Geometric Gaussian Graphs

$$x, y \in [0,1]^d, \qquad g(x,y) = \exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right)$$

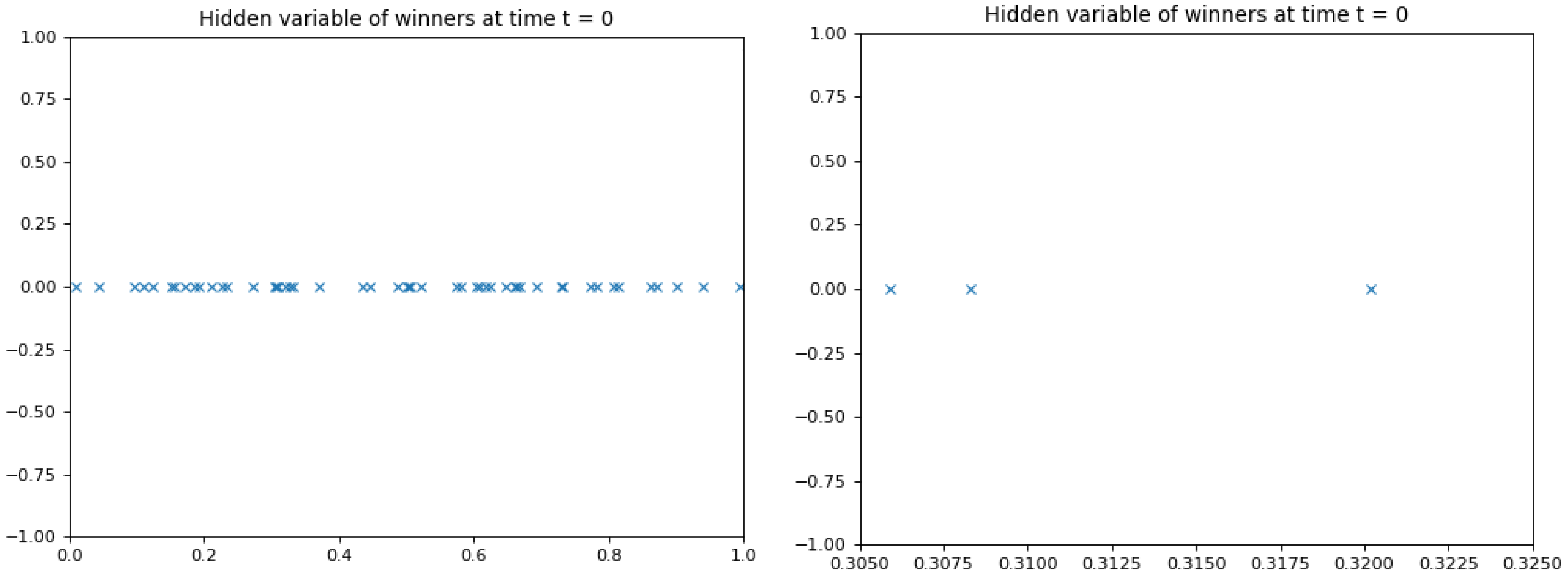
Repeat:

- Next cap $A_{t+1} = k$ vertices with highest degree from A_t.

Theorem[Reid-V.22] With $n > k^{2+d}$ vertices, for a random subset A_0, after $t \geq C \ln^5 k$ steps, WHP A_t will lie in a ball of radius $\sigma\sqrt{\frac{\ln k}{k}}$.

Convergence on Geometric Gaussian Graphs

Theorem[Reid-V.22] In any constant dimension d, with $n \gg k^{2+d}$ vertices, from a random subset A_0, after $t \geq C \ln^c k$ steps, WHP A_t will lie in a ball of radius $\sigma\sqrt{\frac{\ln k}{k}}$ and points in a ball of radius $\frac{\sigma}{\sqrt{k}}$ are chosen nearly uniformly for each subsequent cap.

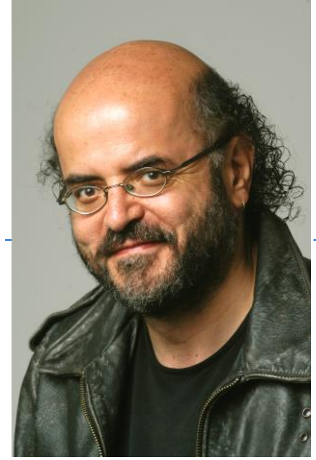


$n = 10000$

$k = 100$

$\sigma = 0.01$

Thanks to



- Christos Papadimitriou
- Dan Mitropolsky, Mike Collins, Wolfgang Maass
- Max Dabagia, Mirabel Reid, Seung Je Jung

… and to you!

Less …. is More?

- Tiny working memory
- Slow, serial processing
- Very little energy usage

- But robustly accomplishes a range of cognitive tasks

- How? What (computational) ideas are we missing?!