Christos Papadimitriou Scribes: K. Gero, Y. Huang, S. Liu, C. DeChant 11/28/2018

Contents

1 Introduction and Recap

Francesco Fumarola (ff2154@columbia.edu) is our guest.

2 Francesco Fumarola' Talk

2.1 Concepts of Free Recall

There is a pyscological experiment in which a subject is made to see or hear a list of words, and then asked questions:

Free Recall: How many words does participant remember after seeing them? *free:* the words recalled need not to be in order.
Recognition: Was this word in the list or not? *answer:* Y/N

2.2 Definition of Variables

M: The number of words presented
Q: The number of words recalled
L: The number of words recognized

Relationship:
$logQ$ and $logL$ is proportional to $logM$.
Mathematically:

$$Q \sim M^{\alpha}, \text{where } 0 < \alpha < 1$$
$$L \sim M^{\gamma}, \text{where } 0 < \gamma < 1$$

Experiments have shown:
$\alpha \sim 0.4$ (Murdock 1960)
$\gamma \sim 0.9$ (Standing 1973)
In this case:

$$Q = M^{\alpha} = (M^{\gamma})^{\frac{\alpha}{\gamma}} = L^{\beta}, \text{ where } \beta = \frac{\alpha}{\gamma} = \frac{0.4}{0.9} < \frac{1}{2}$$

Computer Science goes from algorithm to performance.
Psychology goes from performance to algorithm.
We want to try to model this experiment and come up with these same numbers.

2.3 Fixed points and distribution

Definition: Assume we have N neurons , $i = 1 \cdots N$

$$r_i(t+1) = F(\sum_{j=1} J_{ij} r_i(t) - \theta_i) \text{ , where } r_i(t) \text{ is the firing rate of neuron } i \text{ at time } t$$

Fixed points: (also called attractors, memories, items, patterns, words) In the mathematical field of dynamical systems, an attractor is a set of numerical values toward which a system tends to evolve, for a wide variety of starting conditions of the system. As for free recall, $\vec{r}^{(n)}$, where $n = 1 \cdots L$ and L is the number of words stored in memory (fixed points), obviously $L > M$.

Destabilization: Need to be able to go to multiple nodes, i.e. recall multiple words, and not really get stuck at a single node, so we need some kind of destabilization of the nodes.

Using adaptive threshold θ_i:

$$\theta_i(t+1) = \theta_i(t) - [\frac{\theta_i(t) - \theta_0 - K r_i(t)}{T}]$$

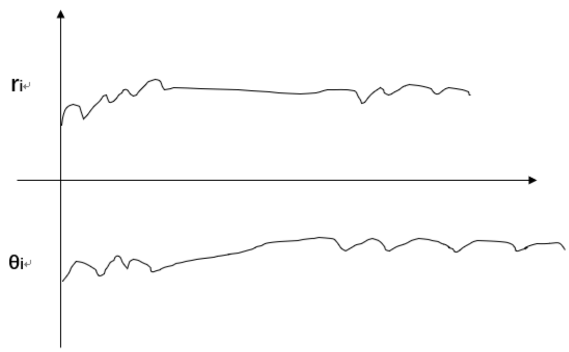

Figure 1: Trend of threshold and firing rate

As threshold increases, firing rate becomes stable and then both threshold and firing rate fluctuates again. The second term is very small, modulated by T.

2.4 Rhos with uniform probability

Think about attractors in brain, how many points are visited before that it goes into the loop (rho)? (Eventually we will visit a point previously visited; then we will start looping.) Suppose we have

$$\phi : V \to V \text{ and } n_{t+1} = \phi(n_t) \text{ then } n_t = \phi^n(n_0)$$

q: The number of nodes in the loop(rho)

P[q;L]: The probability of recalling q stored words out of L

$$Q = < q > = \sum_{q=2}^{L} P(q, L)q = L^{\beta}$$

if $\forall m \neq n, \phi(n) = m \quad with \quad probability \quad \frac{1}{L-1}$, which means the probability of one node to the rest is uniform then

$$P_{closure} = \frac{q-1}{L-1} \text{ and } P_{keepgoing} = 1 - \frac{q-1}{L-1}$$
$$Q = \sum_{q=2}^{L} q\frac{(L-1)!(q-1)}{(L-q)!(L-1)^q} \sim \sqrt{\frac{nL}{2}} \sim L^{\beta}$$

2.5 Rhos without uniform probability

Let's say we have N neurons and $\sigma_i = 0, 1$

For L fixed points: $\sigma_i(+ \to \infty) = \eta_i^n$

$$S_{nk} = \frac{1}{N}\sum_i \eta_i^n \eta_i^k$$

S_{nk} shows how impact of neurons is close to each other.
For example:

$$\begin{aligned} \eta^k &= 0 \quad 0 \quad 1 \quad 1 \quad 0 \quad 1 \\ \eta^n &= 0 \quad 1 \quad 1 \quad 0 \quad 1 \quad 1 \end{aligned}$$

Then $S = \frac{2}{6} = \frac{1}{3}$
$\phi(n) = argmax(S_{nk})$ where $k \neq n$

$$P(\eta_i^n = 1) = b \quad and \quad P(\eta_i^n = 0) = 1 - b$$
$$Q = \sum_q qP(q; L; b) = L^{\beta(b)},\ \beta = \frac{1}{2}\frac{1-b}{1+b}$$

3 Recap from last week

We had two great lectures. The bottom line is that language is fascinating. Phonology, morphology, syntax, semantics, pragmatics. Parsing is hard.

4 Language and the Brain

4.1 History of Langauge

3 MYBP the hominid group separated from the chimps. Only homo sapiens seem to have language. The evidence for lack of language before homo sapiens consists largely of the lack of trappings of symbolic behavior such as figurative art.

80KYBP the first figurative art was made in Africa. Did language come about the same time?

Chomsky (and the French academy in 1866) say that the evolution of language is not a scientific question. Christos wonders whether the human fascination with language is a kind of human speciesism?

There is much speculation about how language came about. Was there a big bang or was it the result of gradual progress? Chomsky and Berwick think that we had language far before 80KYBP, maybe even 500KYBP. Corballis has a different perspective, advocating for the "gestural origin of language" (Wiley 2010).

If we had language in a weak sense before we had what we now call language, what is this "weak sense"? Just hierarchically structured thought? Inner language?

If you have symbols in the mind, you express them in the handiest form possible. and language may not be that handy because it involves a lot of sounds and coordination between people of sounds and their meaning. Perhaps painting or art were easier ways of externalizing these symbols in the mind. Maybe this is why some (e.g. Corbalis, Chomsky) think that art came first. Anthropologists generally think that language is a part of a group of behaviors.

4.2 What is needed for language?

Some think that a (prohibitively) sophisticated muscular system in the mouth, face and tongue is a necessary prerequisite for language. Others disagree and suggest that whatever limitations prevent other primates from speaking is all in the mind.

Corballis thinks we started speaking long after we physically could have. A lingering question is why the language of thought was an evolutionary advantage. Sometimes there is just one mutation that leads to huge effect - e.g. one gene that allows flies to survive in the cold. Language of thought must have been similar.

It is interesting to reflect on the fact that English and Japanese are separated by the changes effected by about four thousand generations of mothers speaking to their children.

If there was a neural big bang, we know where it happened. Our left and right hemispheres are anatomically asymmetric, a condition which seems to be specific to humans. The main locus of asymmetry is the arcuatum fasciculum, which connects Wernicke's area with Broca's area. Long association fibers which connect parts of the brain are much more pronounced on the left than right hemisphere. The arcuatum fasciculum is not myelinated until the second year of a human's life, when he or she is about to form sentences.

The rise of association cortex seems particularly important. Association cortex is cortex that is not involved in sensory and motion.

Incidentally the auditory cortex is near the ear, close to Wernicke's area. The core of the auditory cortex has zones corresponding to different frequencies. The right part discriminates tones, music, while the left involves temporal aspects.

4.3 History of speculation on language and brain

1830s: language must be in the left hemisphere 1870s: Broca and Wernicke discovered their areas and thought they knew the associated pathologies (production for lesions in Broca's area and comprehension for lesions in Wernicke's area). But they got it wrong. Broca's aphasia is lack of syntax and grammar. Wernicke's aphasia corresponds to a deficit in word production in which nonsense words may be produced.

A lot of people think Broca's area used to be some kind of motion planning area in other animals. Perhaps relatedly, problems there can lead to inability to follow directions.

Chomsky - Skinner debate 1960. Skinner was a behaviorist who believed in reinforcement as the key to learning. Influenced by the tabula rasa line of thought, he contended that humans get everything we know from stimuli. Skinner's book "Verbal Behavior" went too far. Chomsky had a devastating critique.

Chomsky theorized that language was innate because he argued that there is an enormous gap between stimulus and competence. He wrote that grammar is innate and universal and that children only have to tune it with 'details'. The extreme minimalist program for understanding the rules of language is that all you need is "merge." Proponents of this believe that there is effectively one grammar rule: there is a noun phrase and then a verb phrase that says what the noun is or does.

Human language is widely considered different from other animals' forms of communication because it is characterized by recursion and infinity. (Note that this is despite the Piraha controvery which errupted when a language spoken in the Amazon was found that apparently does not use recursion (Daniel Everett 2010)).

4.4 The Chomsky Hierarchy

Four increasingly general kinds of (recursive, infinitary) grammars

1. Regular or right-linear: A $\rightarrow$ abB
2. Context free: S $\rightarrow$ aBbS

3 Context sensitive: SB $\rightarrow$ S a b B (NB: the right hand side must be no shorter than left hand side)

4. General: ScD $\rightarrow$ Ba

Right linear grammars correspond to finite state machines or regular expressions (one way reading, finite states).

Context free or phrase structure grammars correspond to nondeterminsitic pushdown automata. (Finite states, one way reading, but can also look at and modify the top of a stack).

Context sensitive grammars or nonshrinking grammars correspond to nondeterministic Turing machines that run in linear space (LBA's). They run only on their input.

Unrestricted grammars correspond to Turing machines that accept by halting.

All recursively enumerable (RE) languages correspond to decision problems 'solved' by Turing machines that reject by never halting.

The Chomsky hierarchy had a tremendous influence on Computer Science, leading to much of the research agenda of Theoretical Computer Science in the 1960s. It helped us understand how to write compilers and trained us for the real problems to come (e.g. P = NP?).

4.5 Language and the Brain

The mirror neuron system helps us understand the actions and intentions of others. It is found in humans and other primates (and there may be something like it in songbirds). It seems closely related to language (and it is physically instantiated in the brain in close proximity to language areas of the brain). For more information, see Corballis and Lakoff who think that language is a corollary of the mirror system that hijacks things from the motor system, which is why they think gestures came before spoken language as a form of communication.

4.5.1 Experimental findings

The Poeppel 2016 Experiment:

Someone speaks words in monotone voice at the natural rate of speaking, 4Hz. Naturally there is synchronized activity in the brain at 4Hz.

Then the speaker makes every 4 words a sentence, e.g. "bad cats eat fish." Then there was a peak at 4Hz but also one at 1 and 2 because once a second you have to hear a word and then you have one at 2 Hz because you have to recognize a two word meaningful chunk (e.g. bad cats) twice each second.

Tree diagrams are being formed in the brain somehow. Many linguistics agree, including Chris Manning.

Frankland and Greene PNAS 2015:

Different areas of the Superior Tegmental Gyrus (STG) respond to "truck" in these two sentences: 1. The ball hit the truck 2. The truck hit the ball.

So the difference between subject and object is localized in the brain - and is done so semantically because sentence with similar meaning but with different structure are processed similarly. Whether a sentence is in the active or passive voice, for example, does not seem affect how it is processed as long as the meanings are the same: "the truck was hit by the ball" leads to similar activity as "the ball hit the truck."

By what mechanism can each tree-building step be carried out by a dozen or so spikes? Recall our discussion of projection and merge over assemblies.

According to Zaccarella and Friederici ("Merge in the Brain") the completion of phrases and especially sentences lights up parts of Broca's area. Broadmann area (BA) 44 lights up for phrases, BA 45 for sentences.

What is the neural gadget that does all this? Consider assemblies of neurons. They seem to encode concepts, presumably also words. They can be projected to other areas. They associate /intersect to reflect affinity. They can create trees by merge.

Here is one proposed brain architecture for syntax. It is the architecture for the plumbing of language, the architecture for generating language, not parsing it. Generating it is more fundamental than parsing it.

When you see a picture first you see an action (e.g. "kick") you search in your lexicon for a verb ... kick. Then you project this to the Wernicke's area location where the verb should be. Then you look again for the noun word, "boy", project it again to the agent part of Wernicke's area. Then you look for the object, "ball", and project to Wernicke's area. Then you hitch a ride through the big fiber, the arcuatum fasciculum, where I propose does the complicated operation, Merge, occurs. Then project to BA 44 (phrases) and BA 45 (sentences).