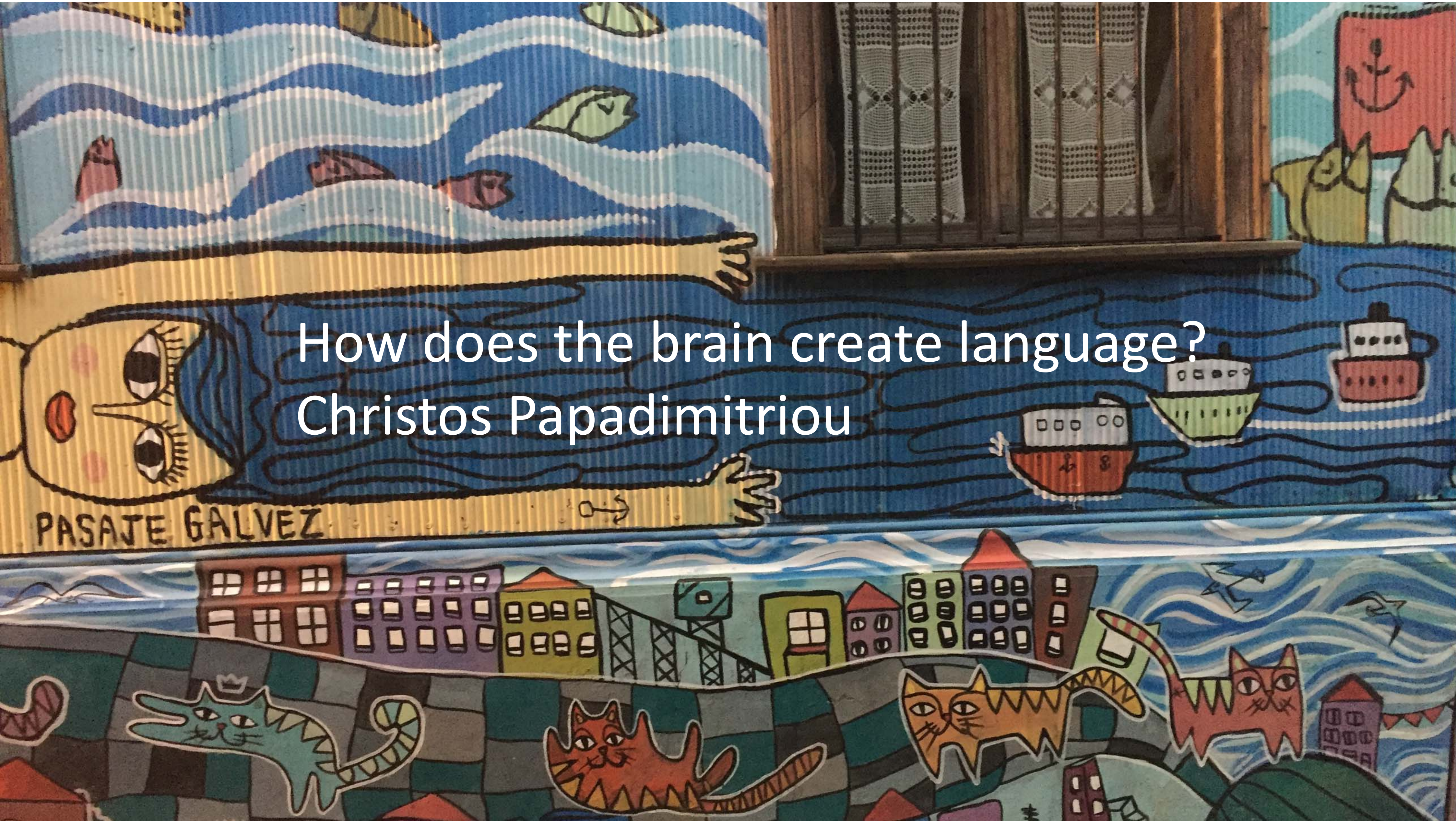
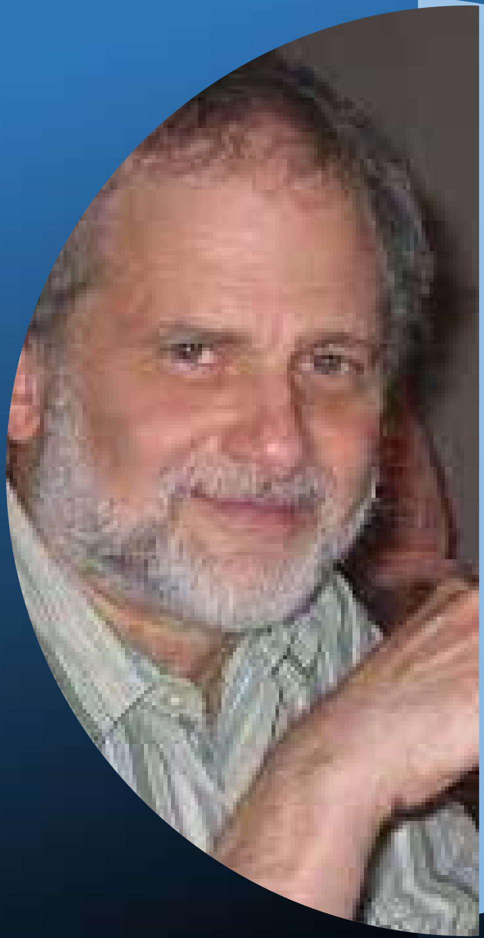
How does the brain create language?

Christos Papadimitriou





my ICASSP connection:
Ken Steiglitz

Fact:
AI still lags far behind brains on several fronts…

- Groundedness
- Inventiveness
- Continual learning
- Emotional and social intelligence
- Energy consumption

An idea: what if we complemented AI with a more ***brain-like, less artificial intelligence?***

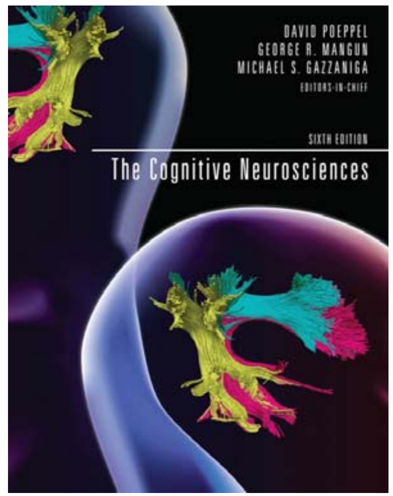
e.g: no backpropagation

[chatGPT, Dec 2022]

How does the brain work?

How does the brain create the mind?

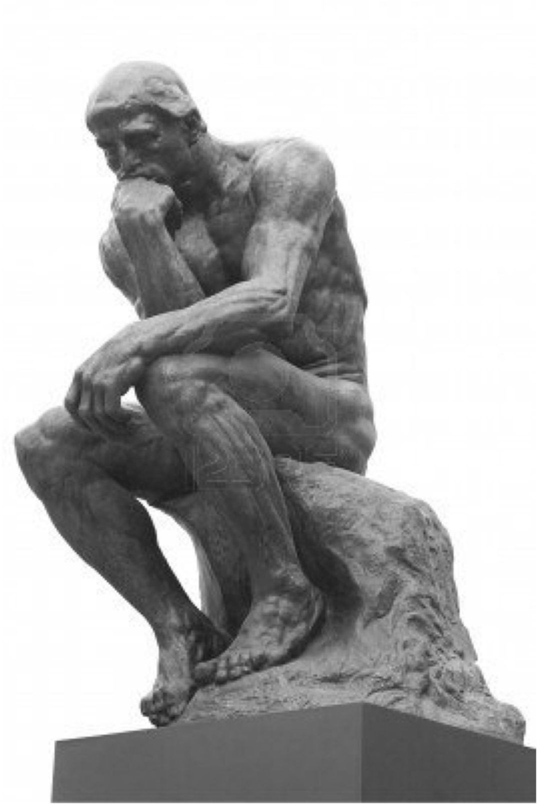
mind the gap!



*“…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

Hmmm, what kind of formal computational theory would qualify as Axel's *logic*?



My approach to Axel's logic

Cognitive phenomena

↑ **implements through simulation**

math model of the brain

↓ **models reasonably well**

Spiking neurons and synapses

NEMO: A math model of the brain

A finite number of brain areas

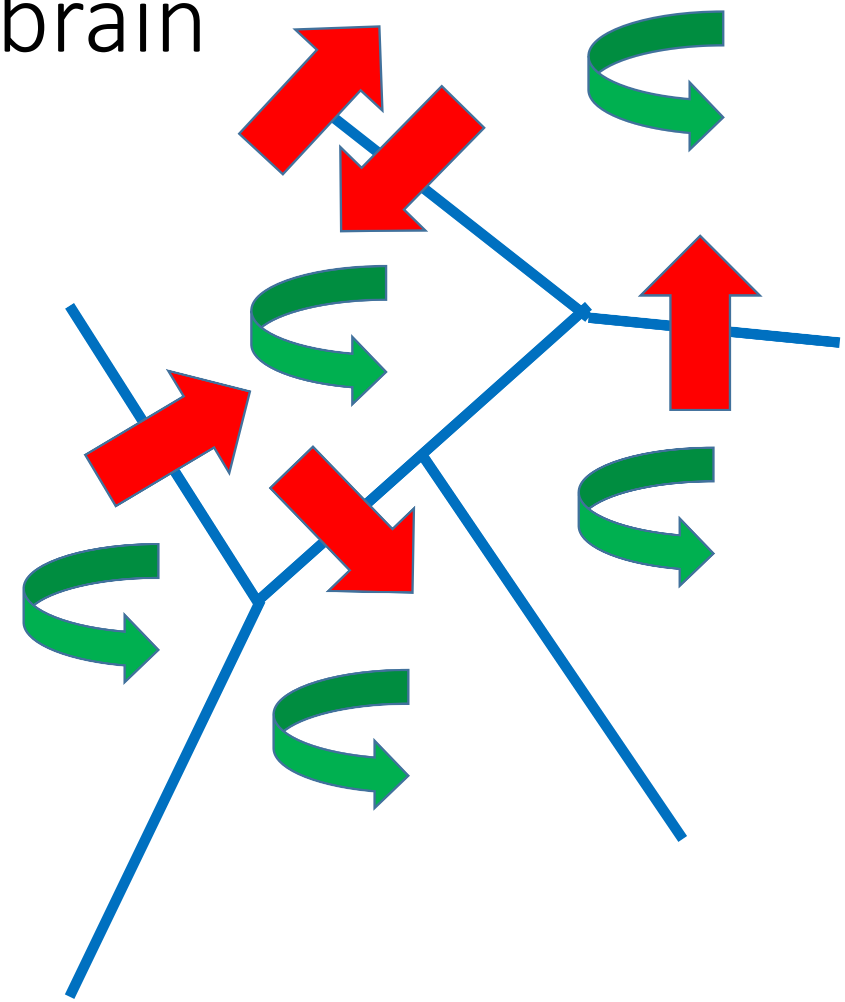
Each has ***n*** excitatory neurons

All areas are recurrently connected

by **random** synapses: $G_{n,p}$

Certain pairs of areas are connected

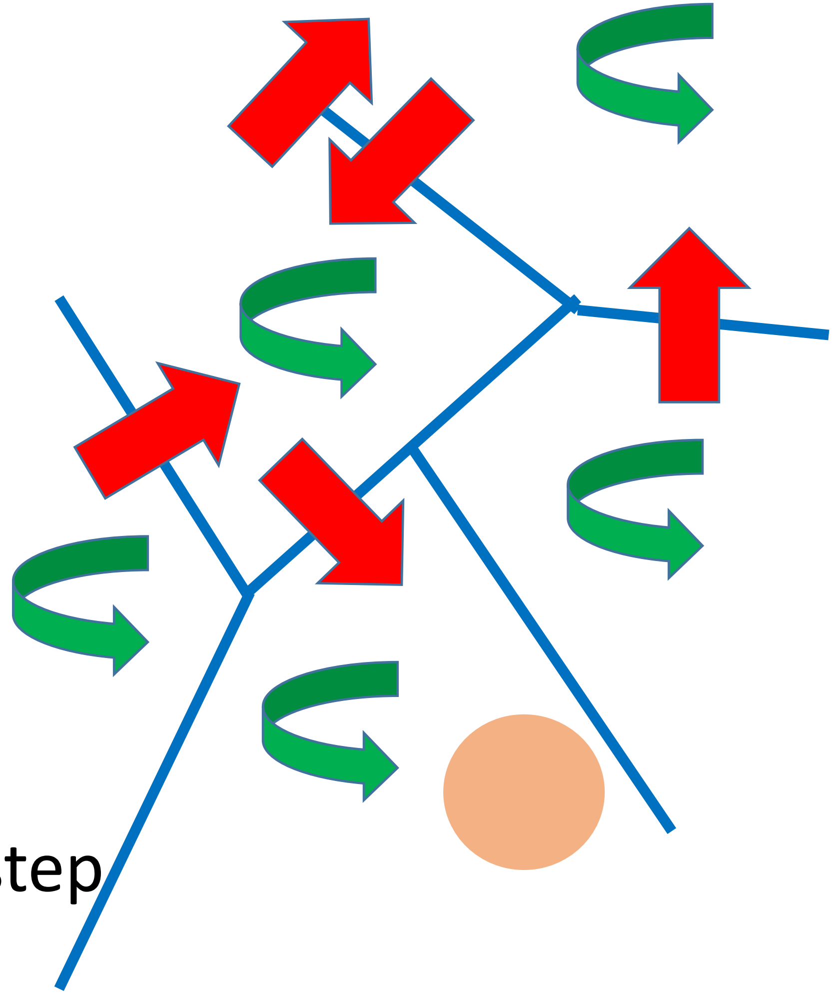
by fibers of **random** connections



NEMO (cont.)

Neurons fire in **discrete steps**
At each step, in each area,
k **<<** ***n*** neurons fire,
those with the highest synaptic input
from the previous step
(models local inhibition)

Also: an area can be **inhibited** during a step



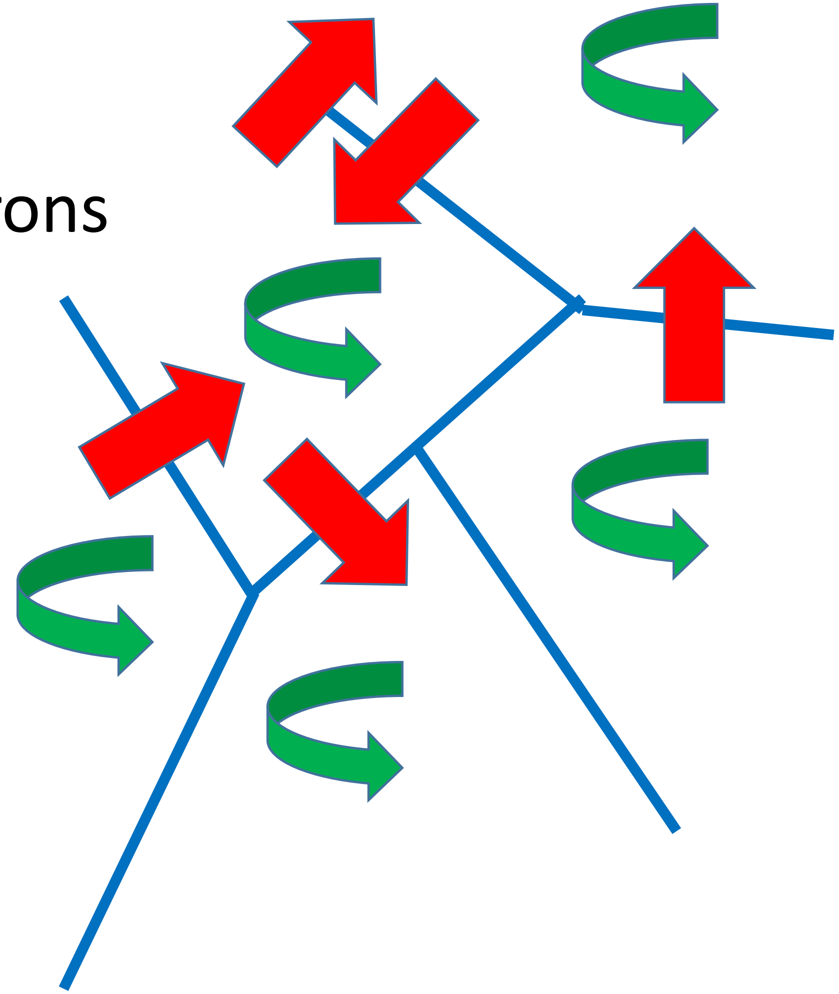
NEMO (cont.)

Finally, **plasticity**: if two connected neurons "fire together" then the weight of their synapse is multiplied by (1 + β)

Plasticity, Randomness, Selection

Typical Parameters:

n ~ 10^{6-7}, k ~ 10^{3-4}, p ~ 10^{-3}, β ~ 5%



NEMO: A neural model

(Also: homeostasis, forgetting)

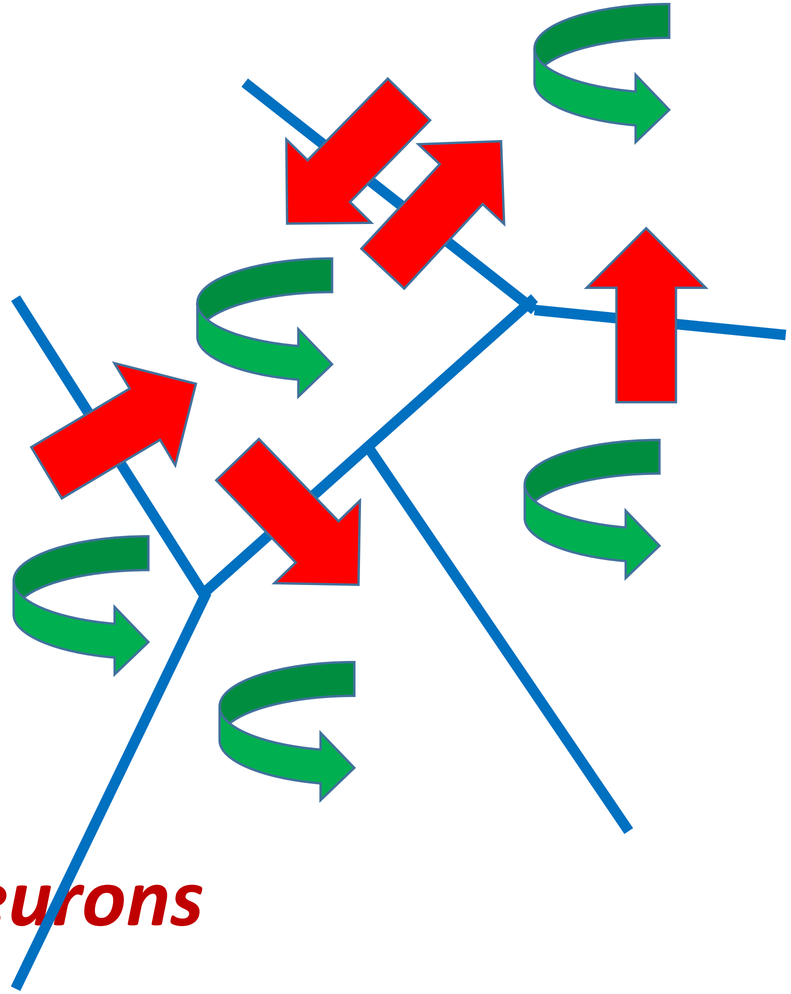
This defines a **dynamical system**

State: neurons that spiked, synaptic weights, inhibited areas

The next step function: fully defined!

Simulator available online

Emergent behavior: ***assemblies of neurons***



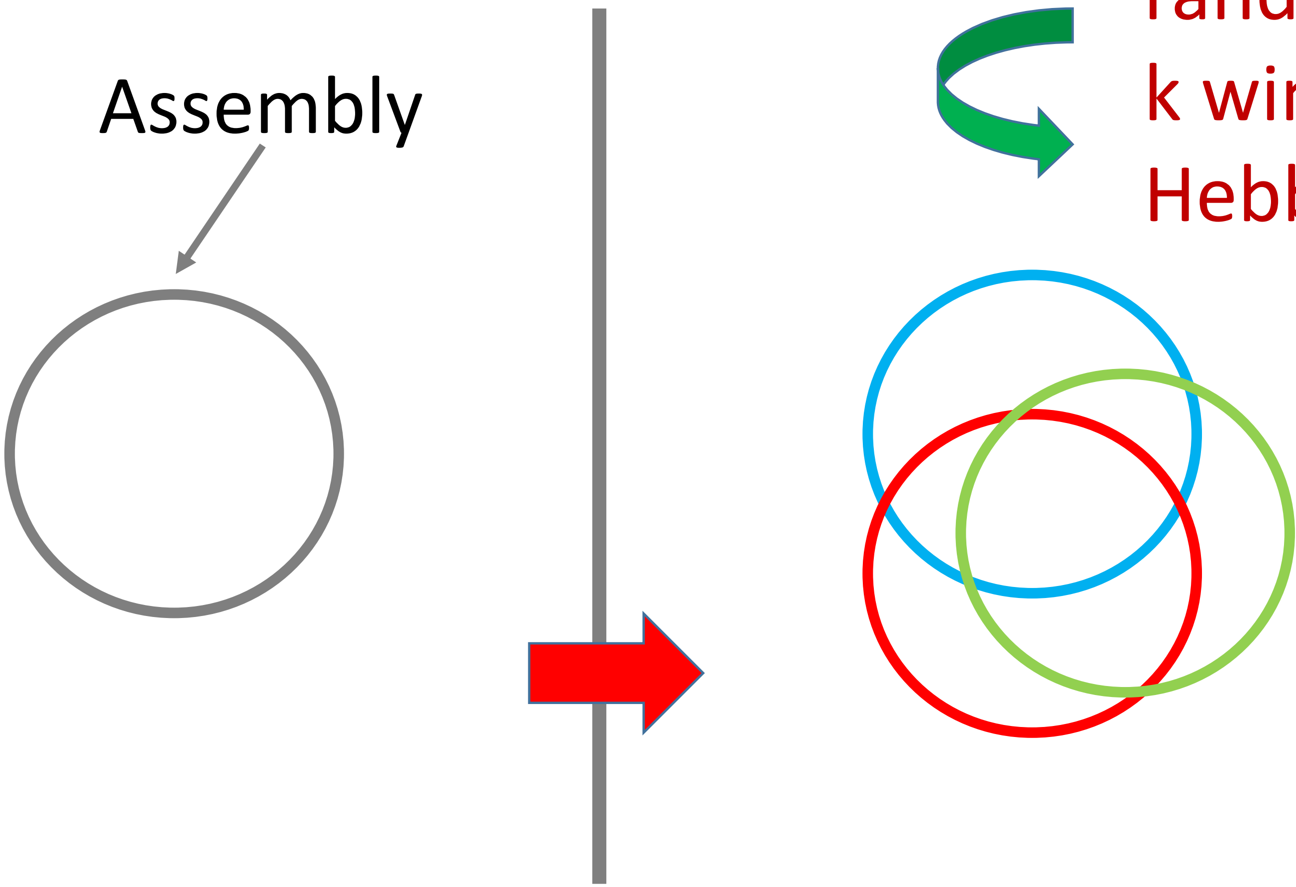
Short history of assemblies of neurons

- [Hebb 1949, Harris 2003, 2005; Buzsaki 2008, 2010, Yuste 2004, 2019,…]
- Assembly: A large ***and densely intraconnected*** set of excitatory neurons in a brain area whose near simultaneous firing is tantamount to the subject's thinking of a particular memory, concept, person, name, word, episode, etc.
- ***G. Buzsaki 2020: "assemblies are the alphabet of the brain"***

Assembly behaviors in NEMO

- Projection
- Reciprocal projection
- Association
- Pattern completion
- Merge
- Sequence recall
- Few shot learning of classification tasks

eg, projection



spiking neurons,
random synapses,
k winners take all,
Hebbian plasticity

Assembly behaviors (cont.)

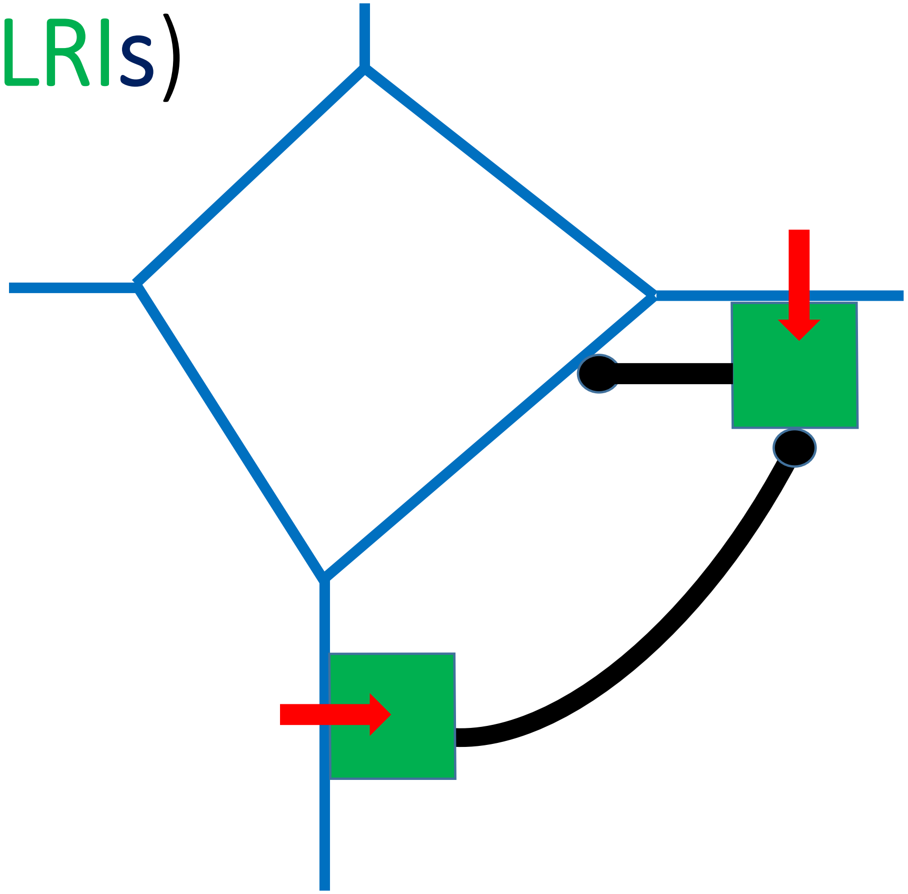
- Projection
- Reciprocal projection
- Association
- Pattern completion
- Merge
- Sequence recall
- Few shot learning of classification tasks

***Theorems** in the math model state that these behaviors function as specified "with high probability"*

*Plus, verified by **simulations***

Q: How are areas inhibited/disinhibited?
A: long range interneurons (LRIs)

- Populations of neurons with long axons
- Extrinsic to the areas
- They can ***inhibit*** remote brain areas…
- …or other LRIs
- They can be **recruited** by the assemblies of an area
- **LRI**s seem to be ***necessary for brain computation*** [Roux and Buzsaki 2015]
- ***They constitute the program of the computation***



assemblies + LRIs = universal computation

A hardware language that can do arbitrary $\sqrt{n/k}$-bit computation

NEMO as software-implemented neuromorphic computation

- To simulate T steps of the model, one needs **$O(pn^2T)$** computer time
- (the number of areas is hidden in the O-notation)
- $T \approx$ #seconds X 50
- A ***lazy simulation*** technique reduces this to **$O(pk^2T^2)$**
- 10^4 speedup, allows us to simulate ***a few seconds*** of brain time on a laptop
- ***Enough for some cool cognitive phenomena!***

Recall our plan: Computation in the brain

- Cognitive phenomena

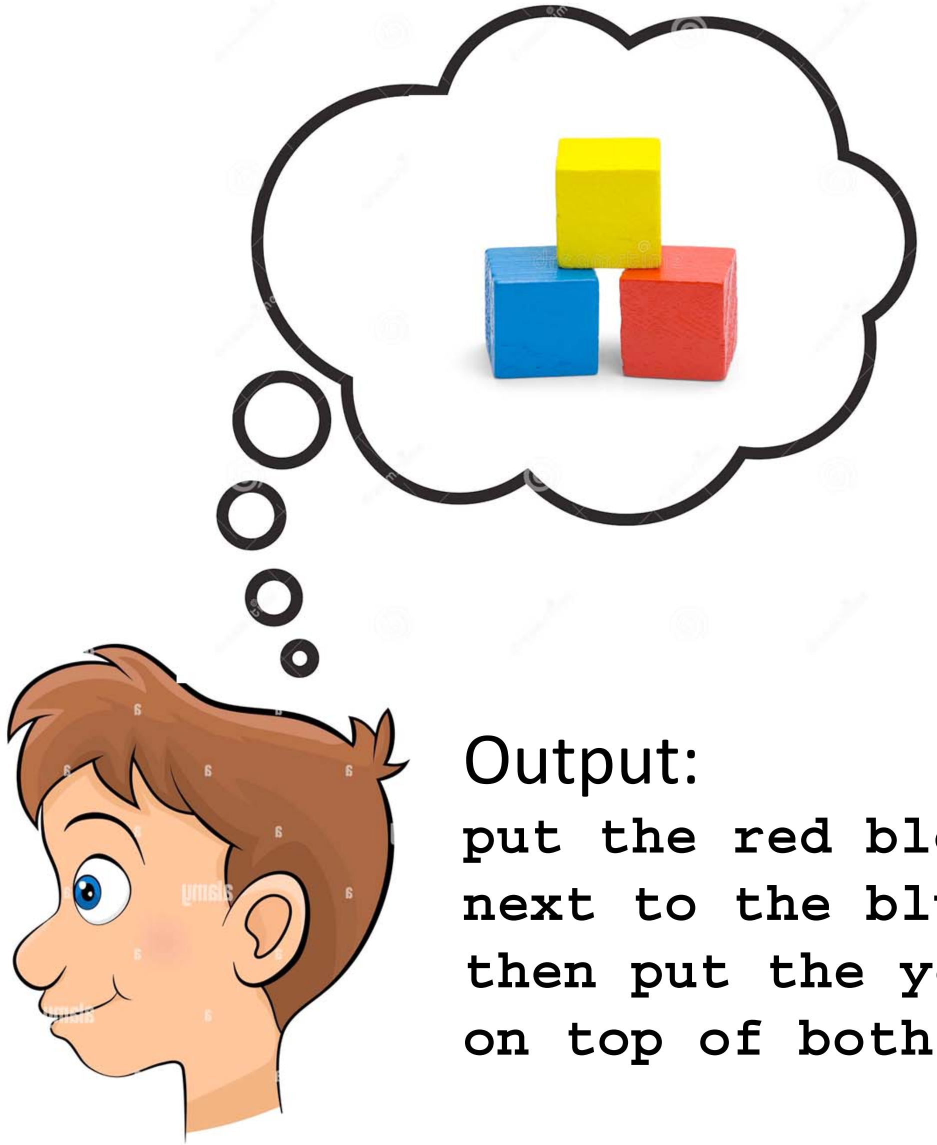
↑ implements *next*

NEMO

↓ models reasonably well

spiking neurons and synapses

Planning in the Blocks World with NEMO (AAAI 2021)

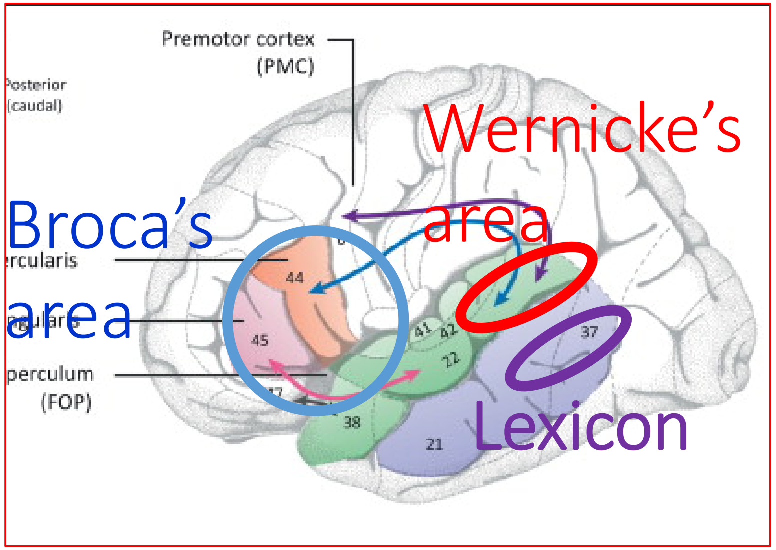


Output:

put the red block on the table next to the blue block; then put the yellow on top of both

OK, what else can we do with NEMO?

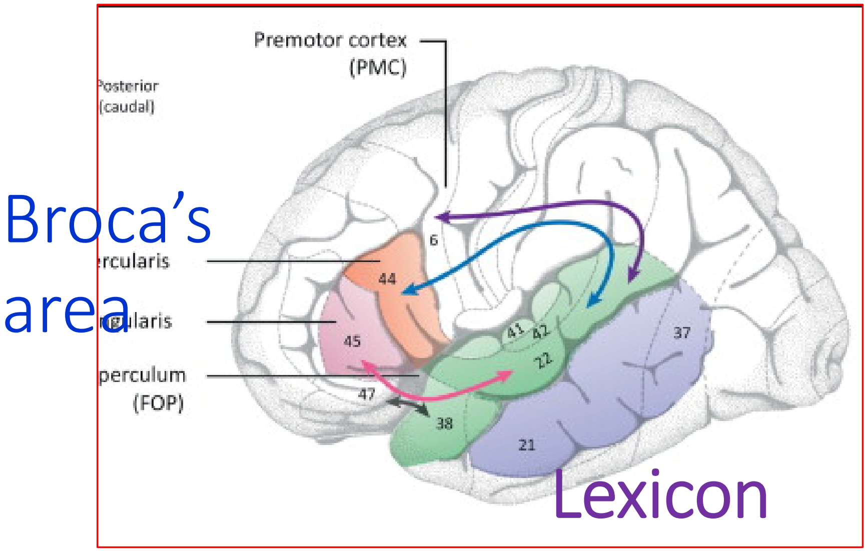
Language!



It is the hardest thing any brain has done

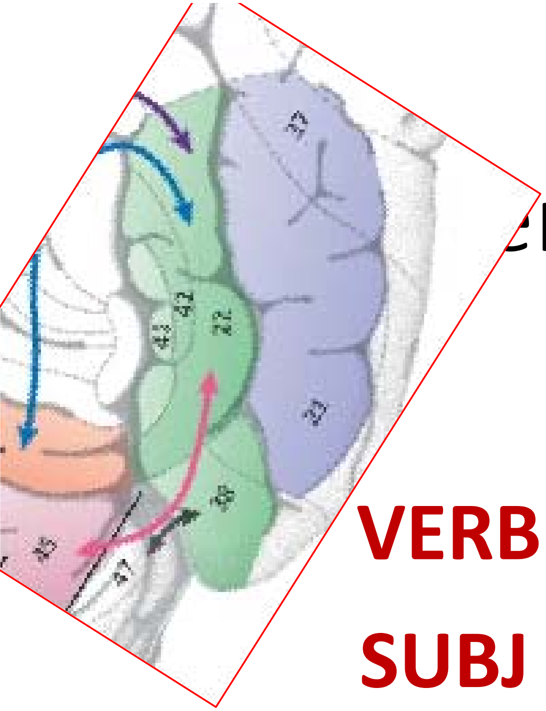
It has evolved over the past 3000 generations, and so it must be a mirror of the human brain

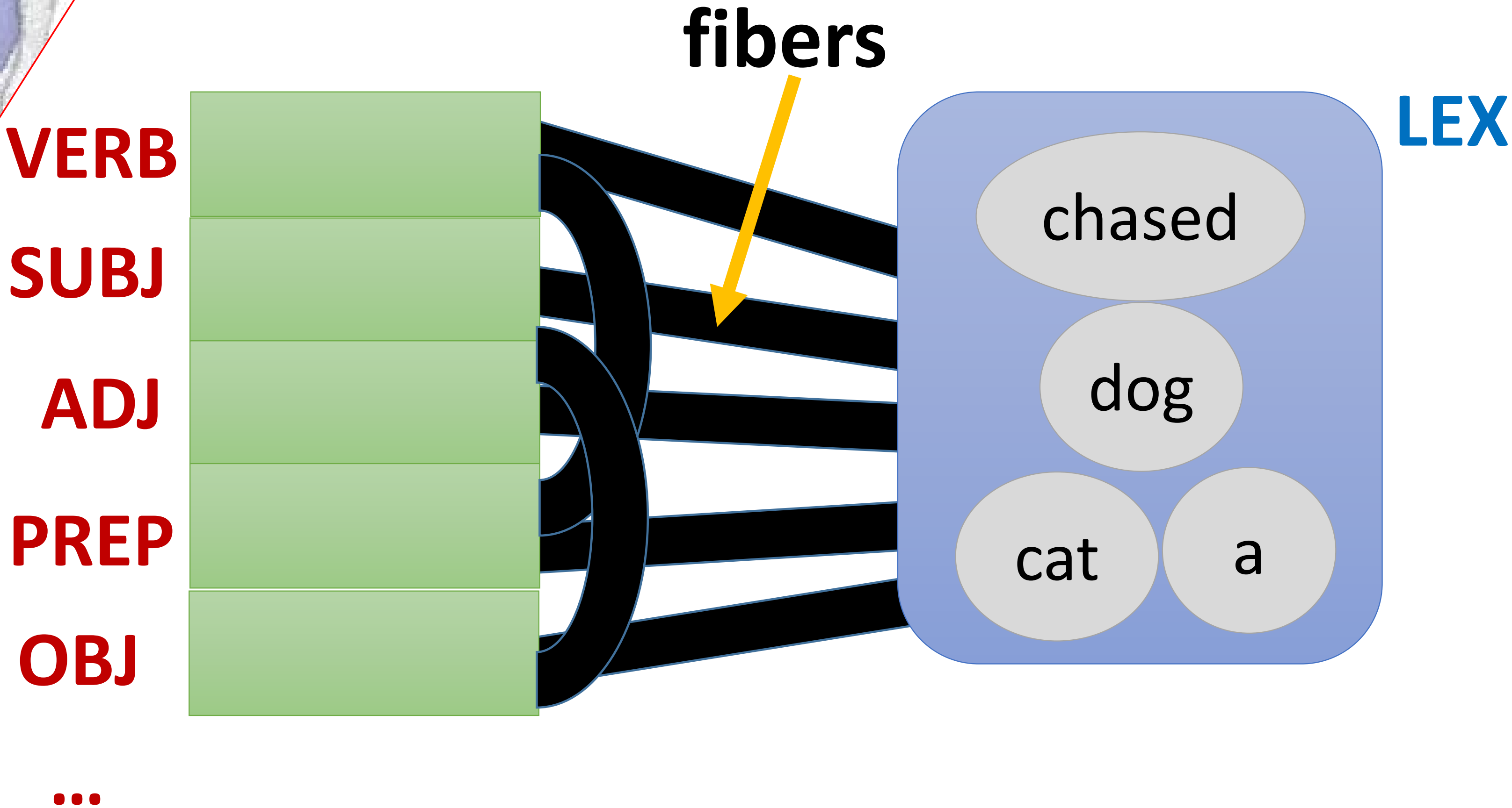
Language



a parser of English implemented by spiking neurons

er: The architecture





The Parser

- Input: a sequence of excitations of word assemblies
- (We assume that **phonology** *has been solved*)
- Each word assembly has an **action set** (inhibit/disinhibit actions implemented by **LRI**s)
- Encodes its ***part of speech***, its syntactic role
- When word assembly fires, its action set is ***executed***
- The sum total of the word action sets ≈ ***the grammar***

Testing the Parser

- Q: what does it mean for this device to *parse* the input sentence?

- A: correct *depedency tree* can be recovered as a side effect

The Parser (cont.)

- It parses simple sentences such as:

"the young couple in the next house saw the old little white car of the main suspect quite clearly"

- Speed: about ***20-25 spikes*** (0.4-0.5 sec of brain time) ***per word***
- Also: simple parsers of ***Russian, Japanese, Hungarian***, ***Mandarin***
- [ACL 2022], code available online

The Parser as neuromorphic computation

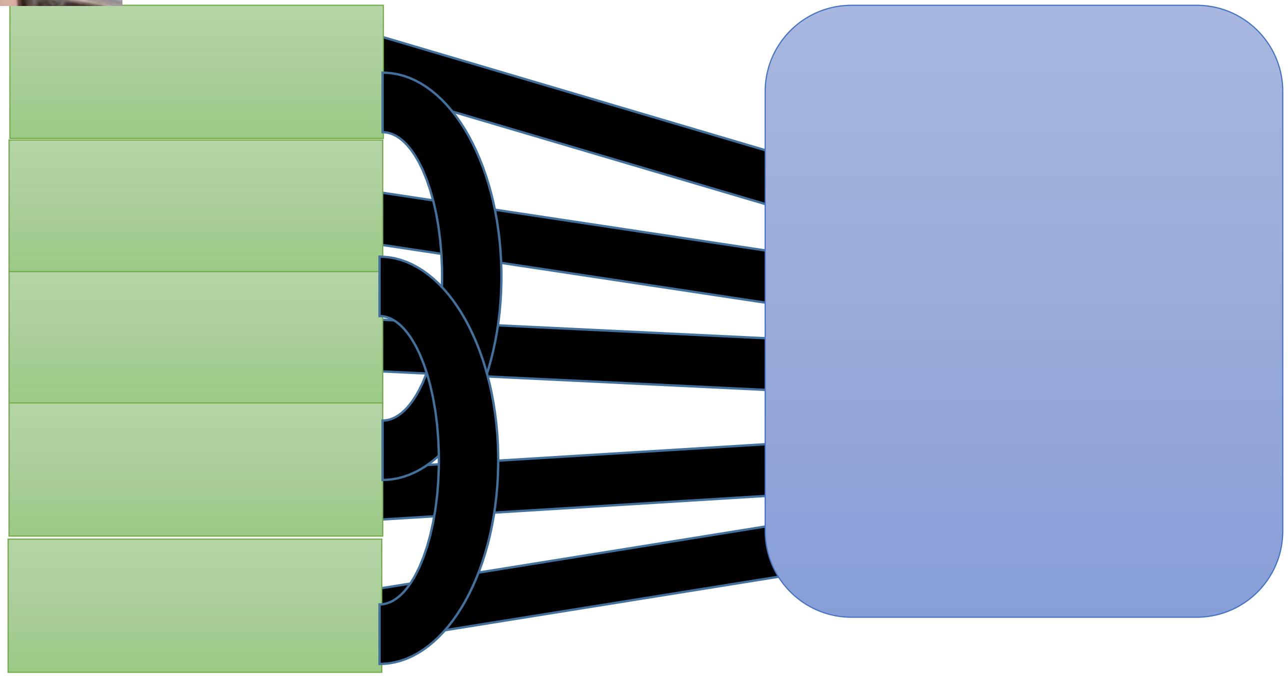
- ***Implemented exclusively through the spikes of stylized neurons***
- 10^7 neurons, trillions of synapses (cf Intel's Loihi2)

Parser FAQ

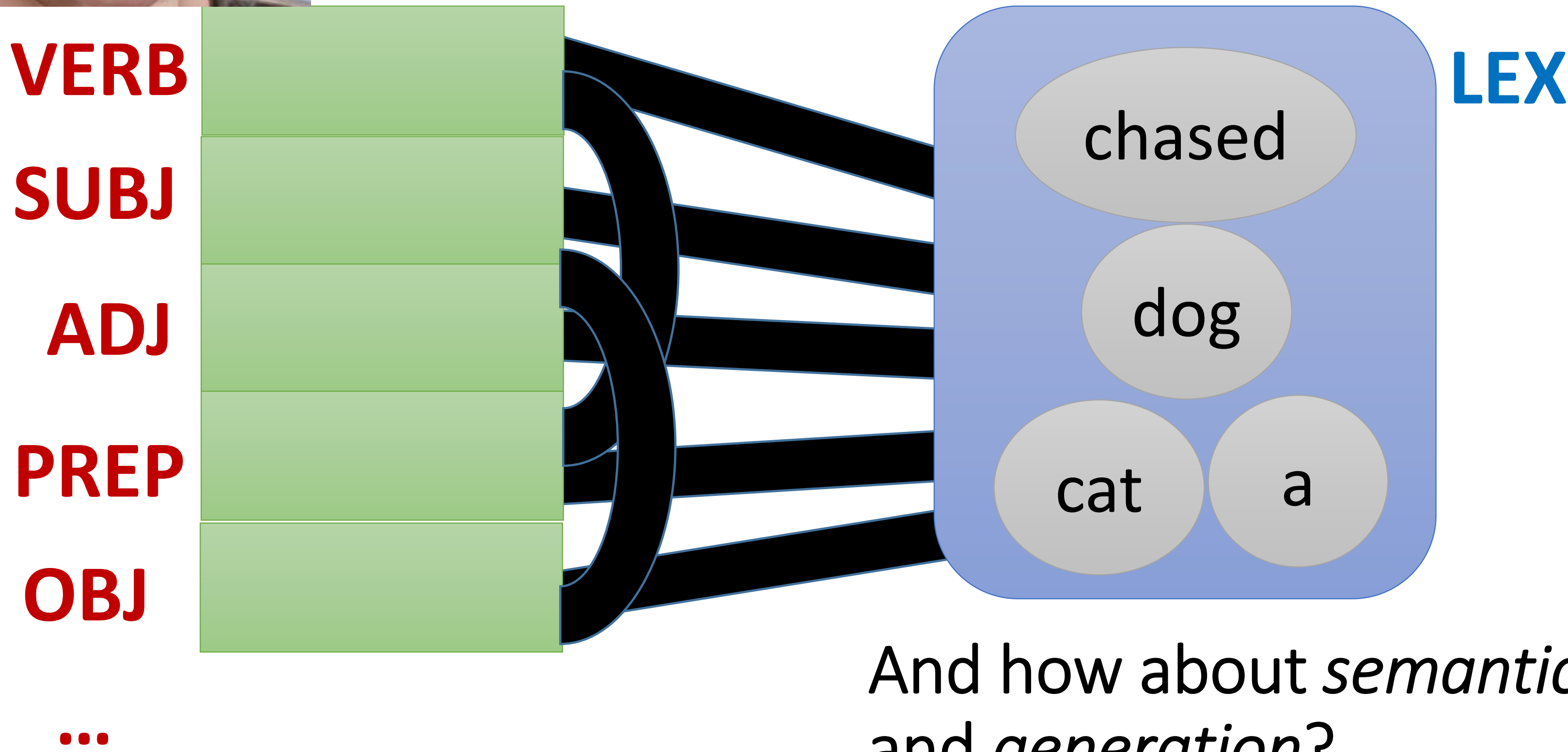
Q: But now we have parsers that can do so much more…

- A: True, but the point of this one is that it is biologically plausible, implemented exclusively through the spiking of stylized neurons, without backprop. It is the first of its kind, and we see no fundamental limits to how good we can make it
- Q: How are you going to test your theory on humans?
- A: We will not. It's not a theory, it's a proof of concept
- Q: But how does this complicated mechanism come about in the infant brain?
- A: Fair point, let's talk about it

How do babies fill the blanks?

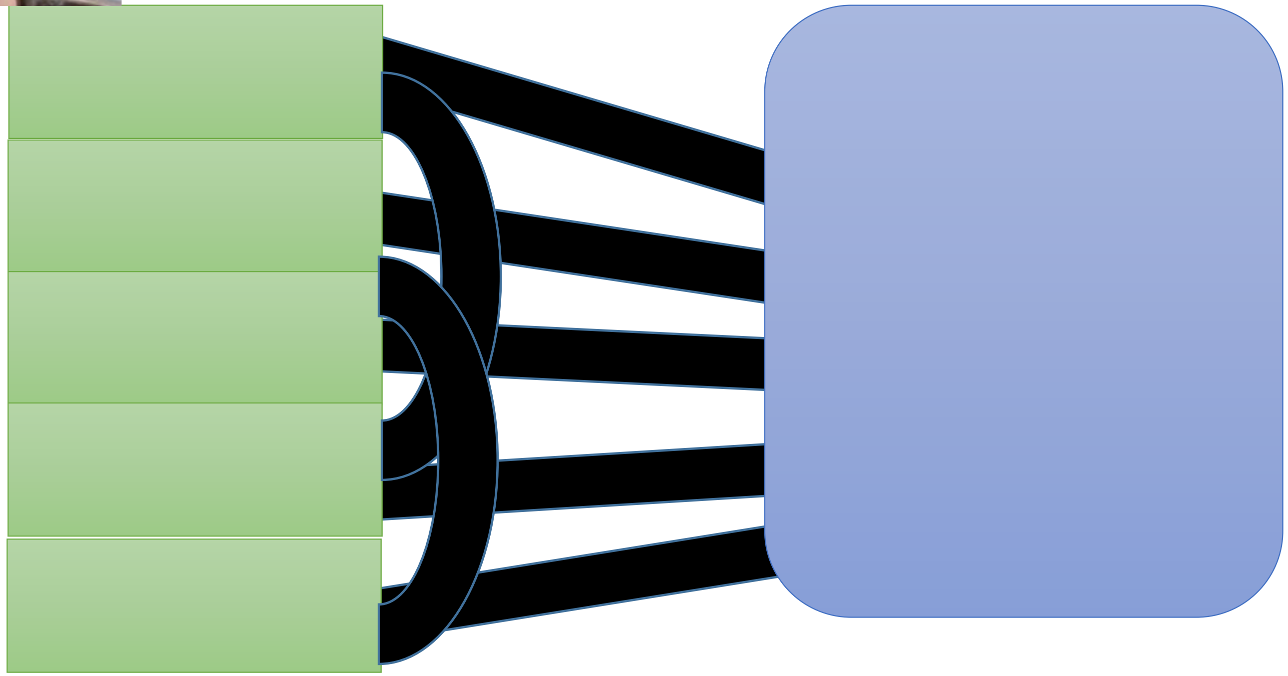


How do babies fill the blanks?



And how about *semantics* and *generation*?

Next: Biologically plausible
language acquisition



Biologically plausible language acquisition

A neural tabula rasa which, upon input of a ***modest*** amount of ***grounded*** language ***with shared attention***, will learn to ***comprehend*** and ***generate*** sentences in the same language

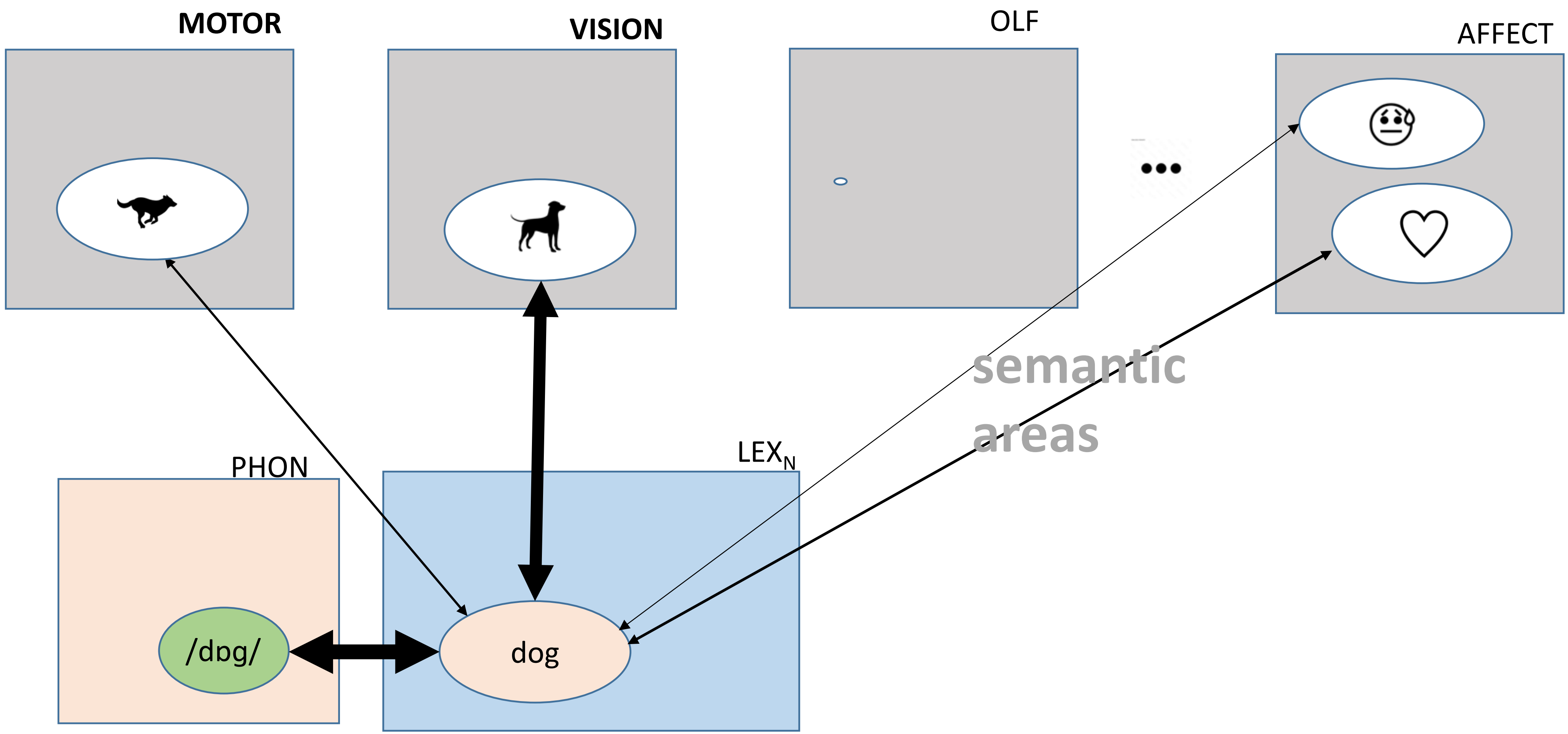
- Modest: O(|Lexicon|)
- Grounded: In the context of the real world
- Shared attention: Tutor and Pupil

Biologically plausible language acquisition: *Representation of words*

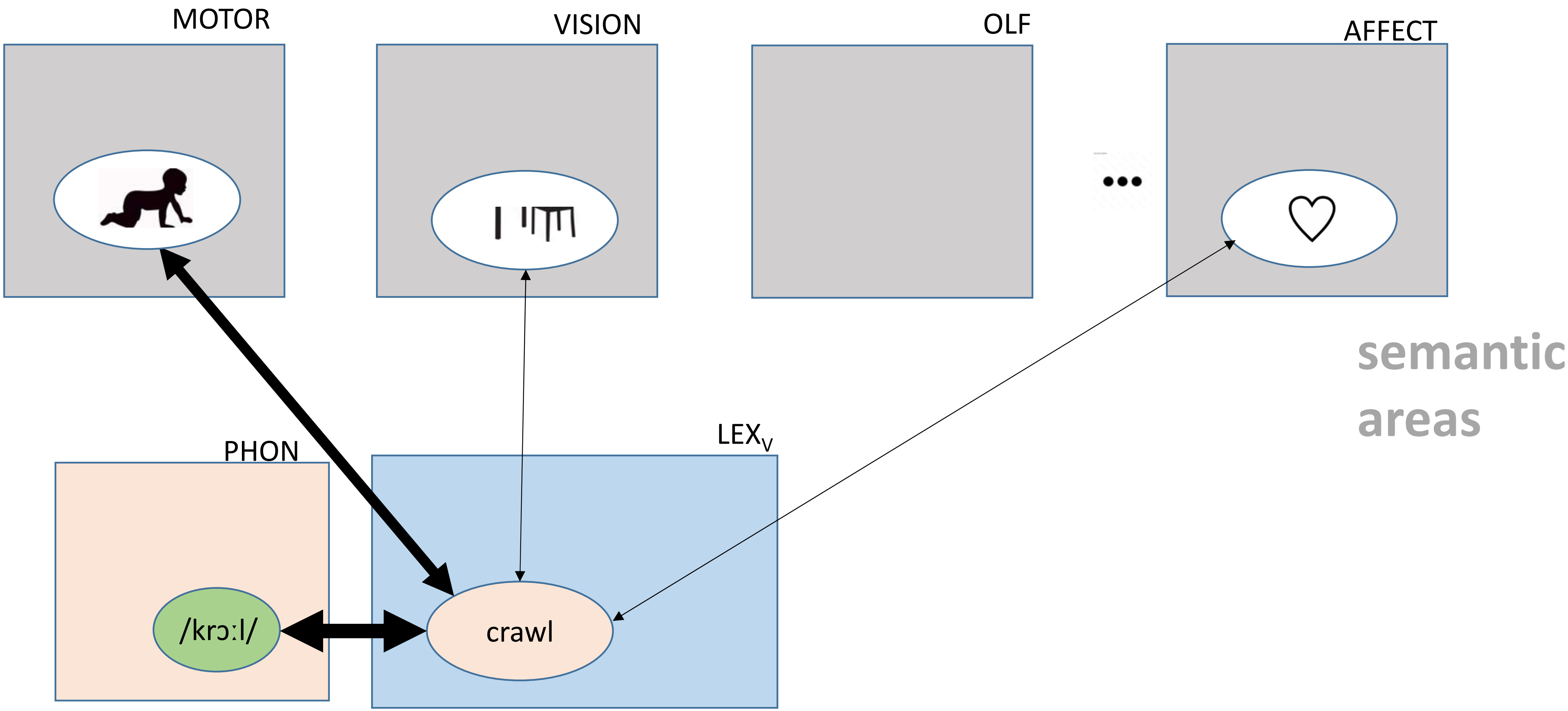
What we ”know” from Neurolinguistics:

- Noun and Verb representations in LEX_N and LEX_V
- ***Phonetic*** representations are separate
- Each word representation is the ***hub*** of a distributed representation of its ***meaning***
- In our system, we have several ***semantic areas***
- Nouns have a basic VISUAL representation, Verbs a basic MOTOR representation

Nouns



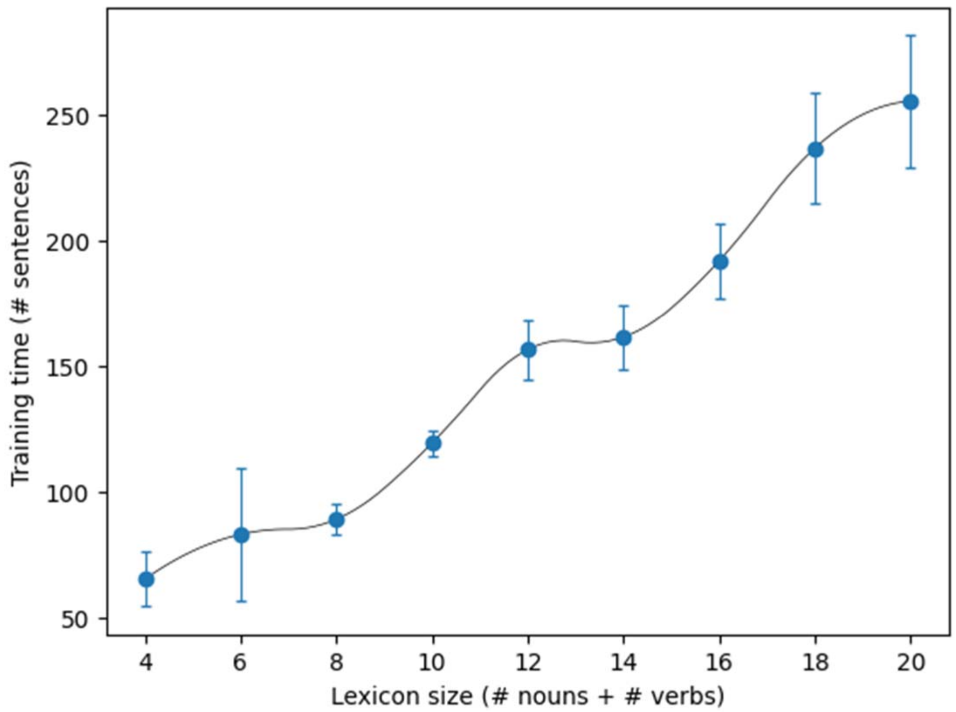
Representation of verbs



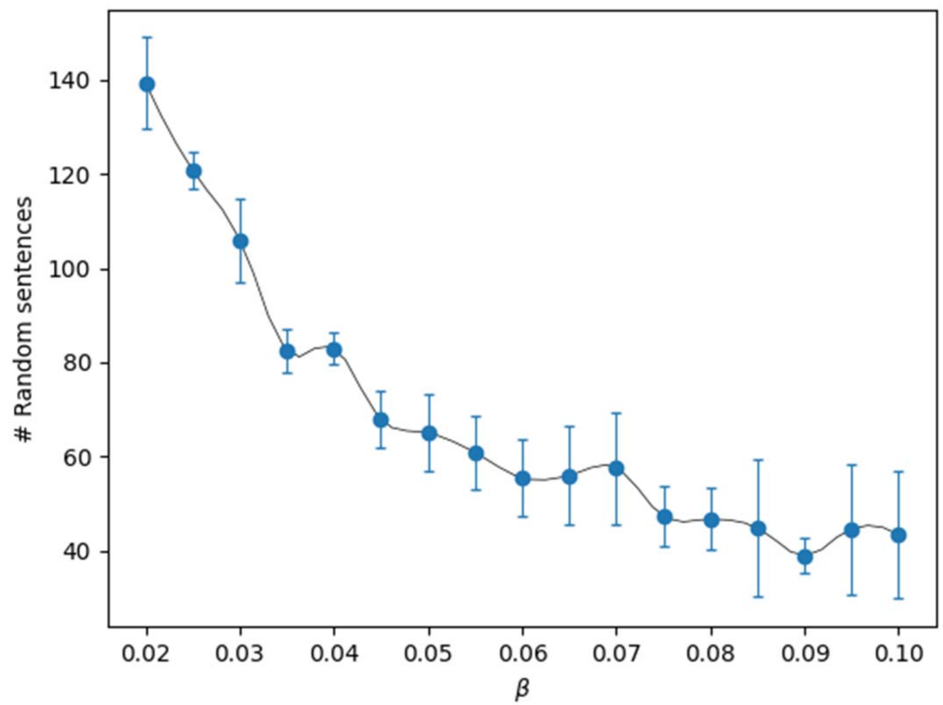
Language acquisition experiment

- For now, only ***intransitive*** verbs: two-word sentences
- We solve the hard case, where the Pupil learns by listening to grounded ***sentences***, e.g. *"doggie runs"*
- Must create, for each word, a representation ***in the correct area***, and ***not in the other***
- Difficulty: *is this **English** or **Irish**?*

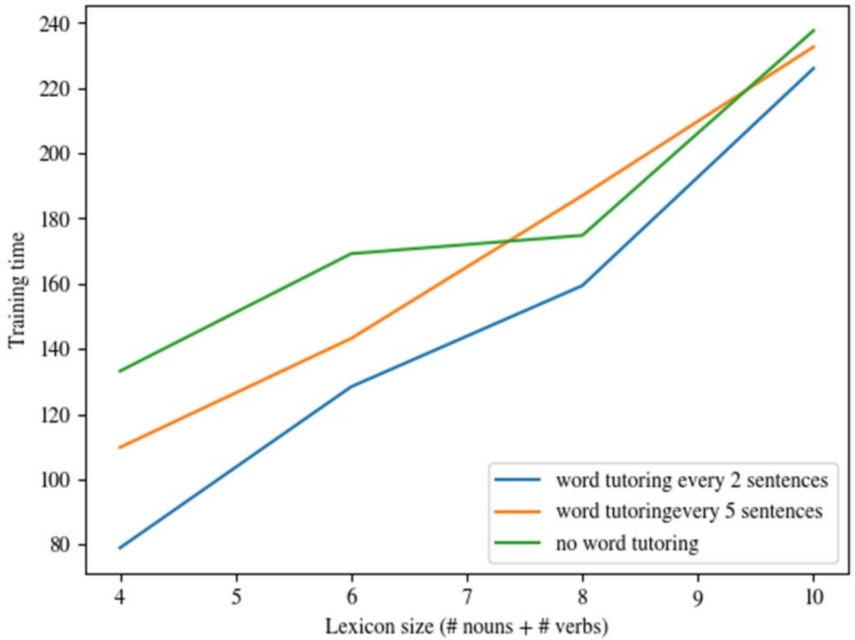
Experiments: how much input is needed



Experiments: as a function of plasticity



Experiments: single-word tutoring helps



Representation of word semantics (cont.)

- The same representation supports ***generation***
- When a cat is jumping (in the perceived world or in the Pupil's imagination), and the Pupil has the ***impulse to articulate***, *"kitty jumps"* will be spoken
- Ergo, the Pupil has learned the language's ***word order***
- PS: ***multilinguality***

Representation of word semantics

- ***A star of assemblies***
- The assemblies of the star, their size and overlap, the strengths of their synaptic connections to the symbol, are all ***dynamic*****,** and encode the ***meaning*** of this word for this speaker and at this moment
- They reflect cooccurrences of words, statistical regularities, order statistics, and changing world contexts

Biologically plausible language acquisition: The next challenges

- Scale to 100 words
- Transitive Verbs and Object category
- Functional words ("the", "she")
- Abstract words (”peace", "contemplate")
- *Subtleties of syntax*
- Develop a suite of brain-like learning mechanisms: one- and few-shot learning, assimilation of patterns and statistical regularities, …

Sooooooo…

- The study of the brain is fascinating and bottomless
- We have just began!
- Implementing language in the brain: close encounters with profound questions
- Progress in biologically plausible acquisition
- **Is NEMO the seat of Axel's "logic"?**
- **Is the design of neuromorphic intelligent systems a useful complementary path to AI?**

my collaborators!

Santosh
Vempala
GaTech

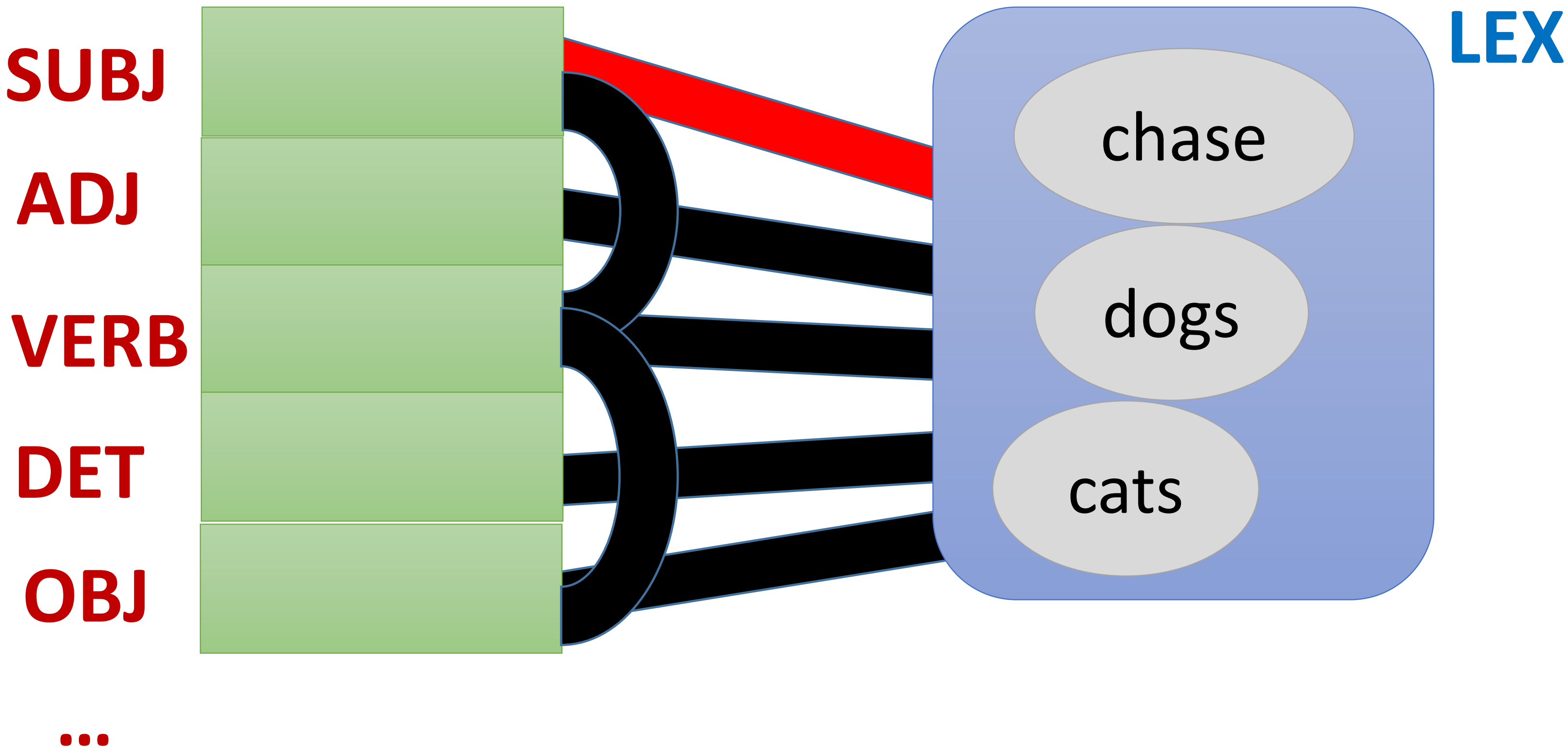
Max Dabagia
GaTech

Mike Collins
Google

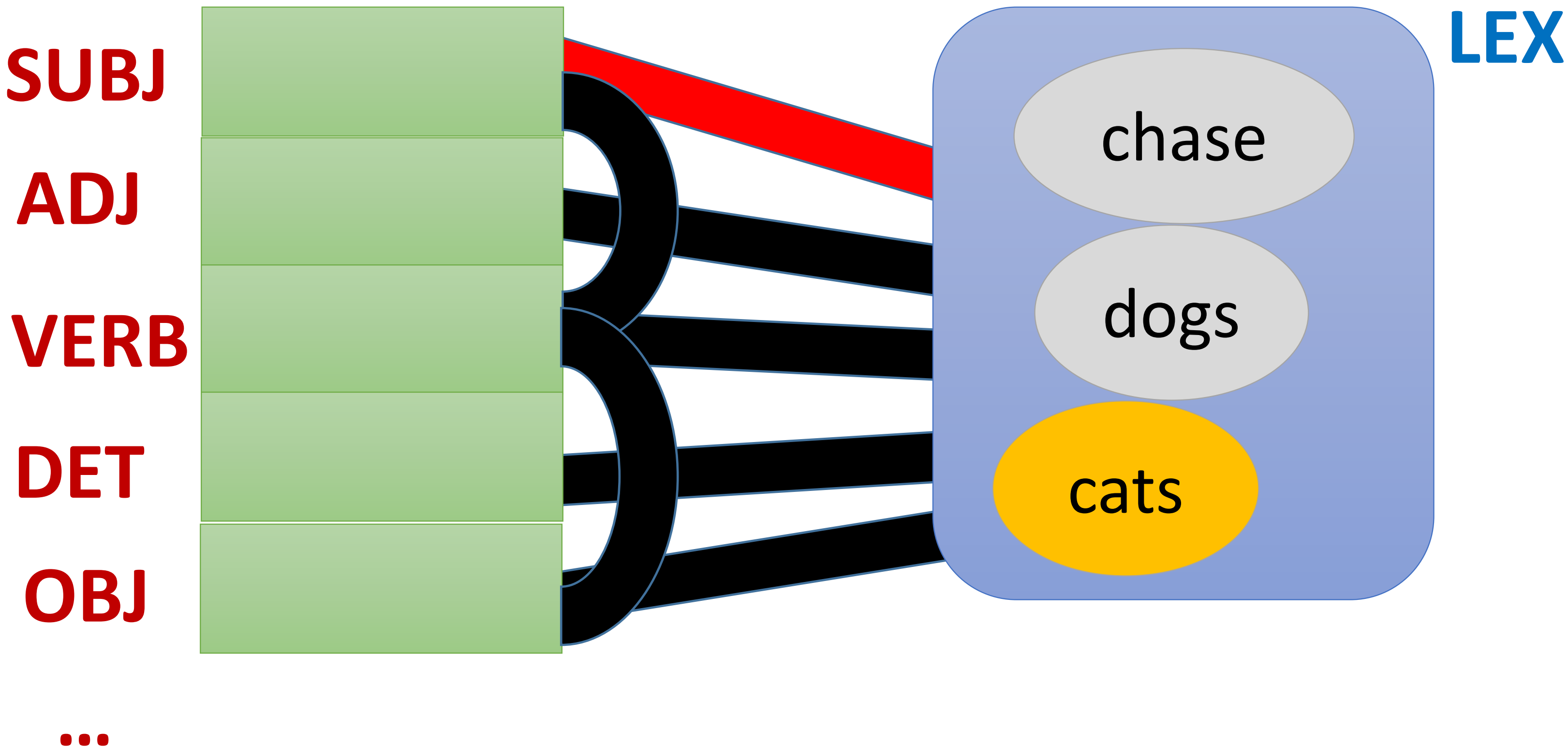
Dan
Mitropolsky
Columbia


ευχαριστώ!
PASAJE GALVEZ

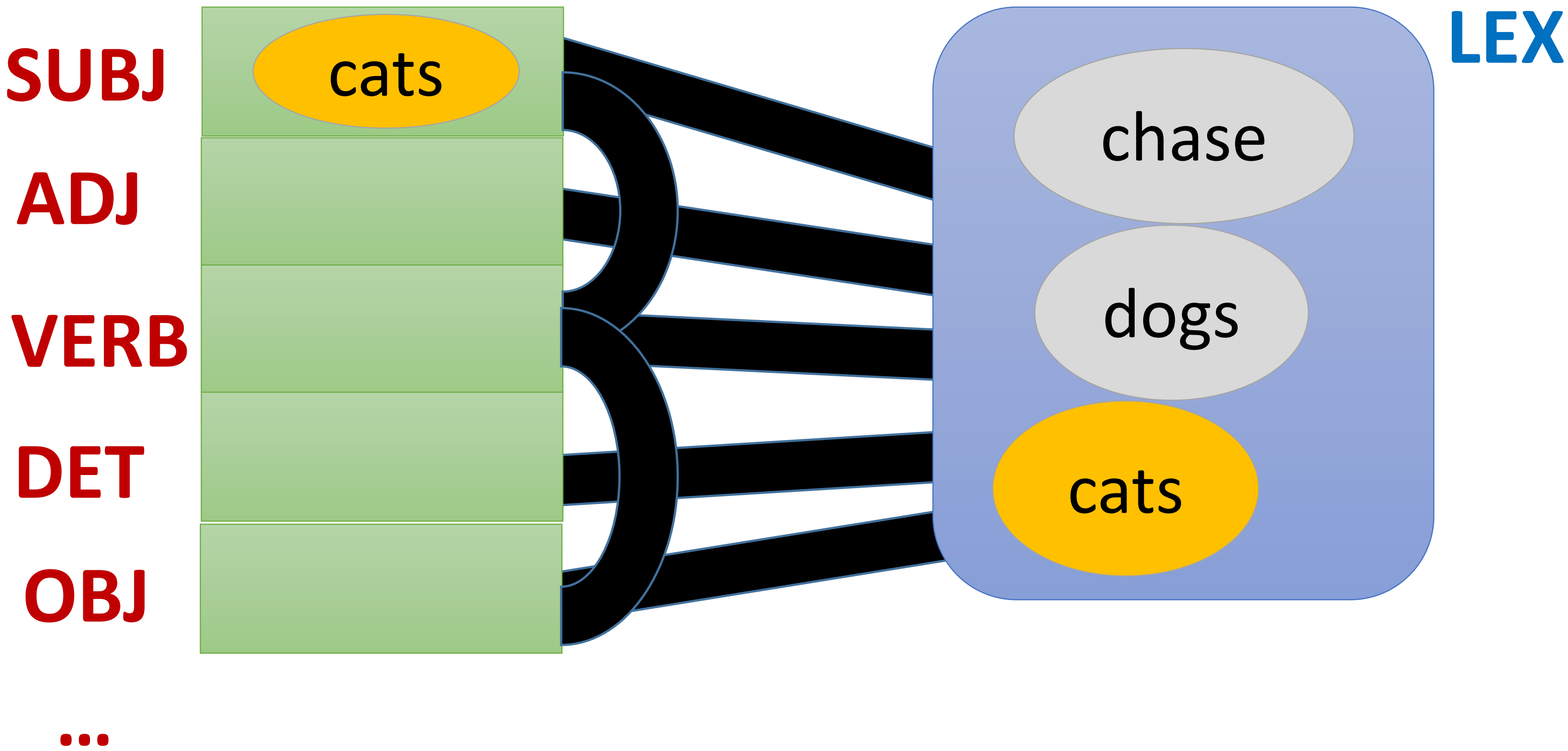
On input **cats chase dogs**



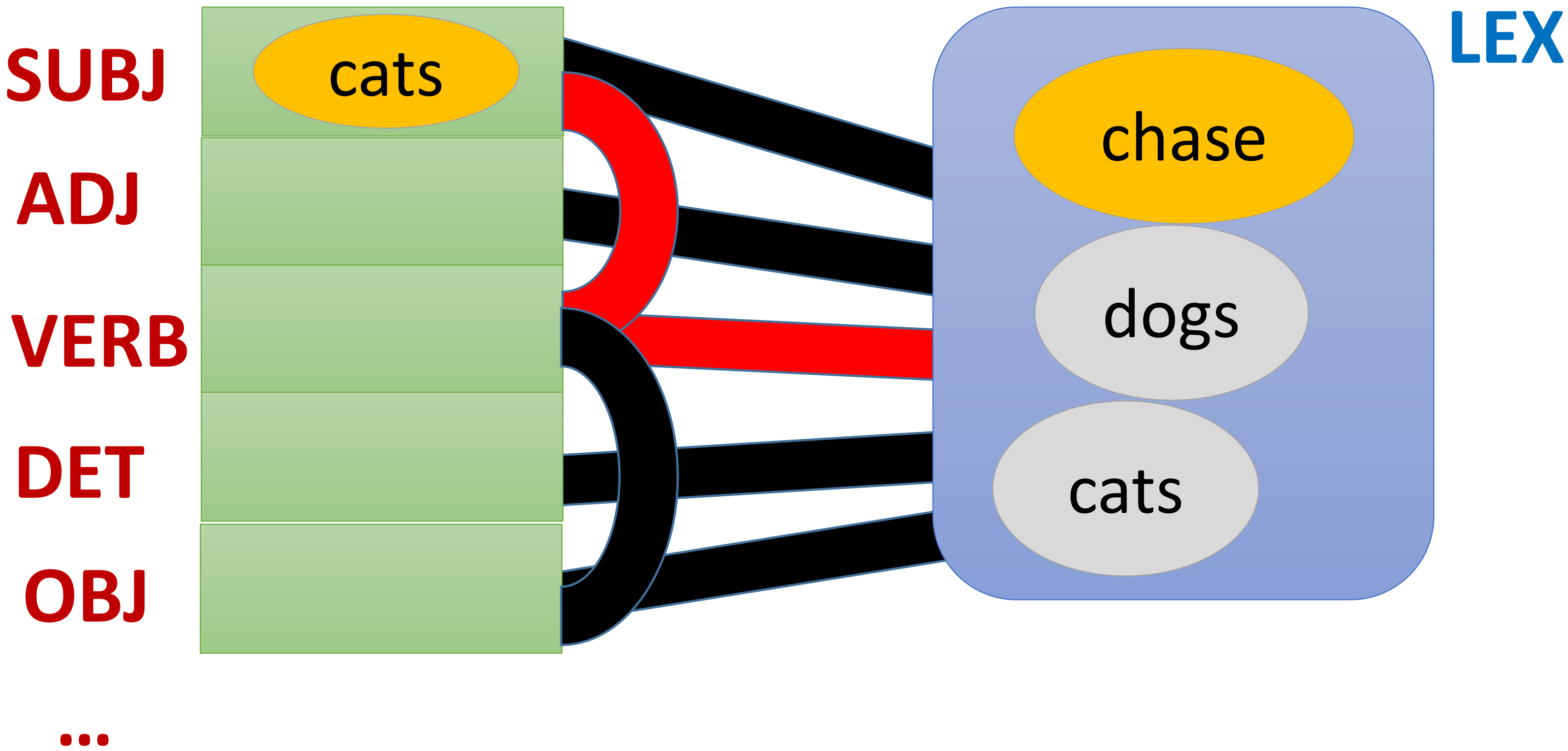
On input cats chase dogs



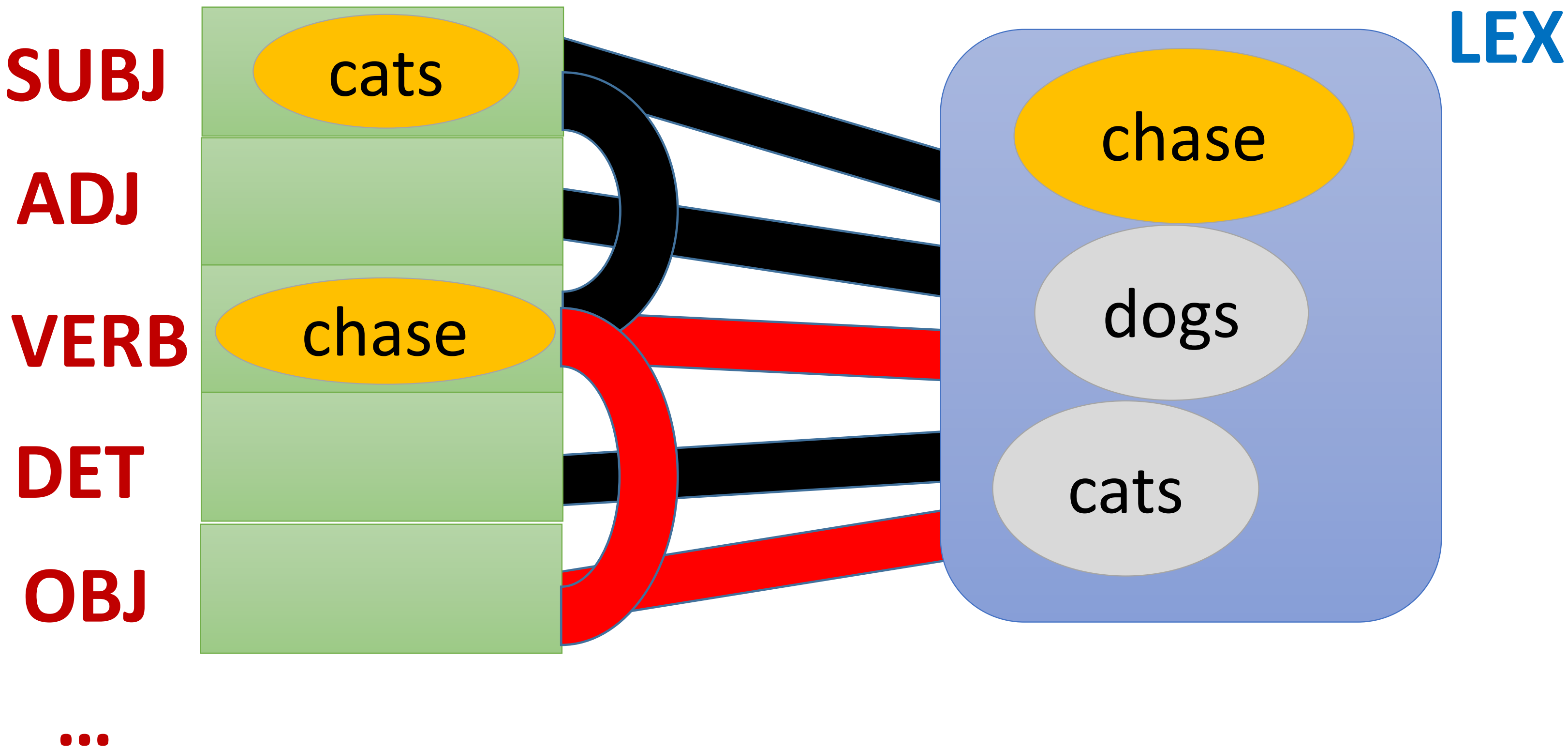
On input cats chase dogs



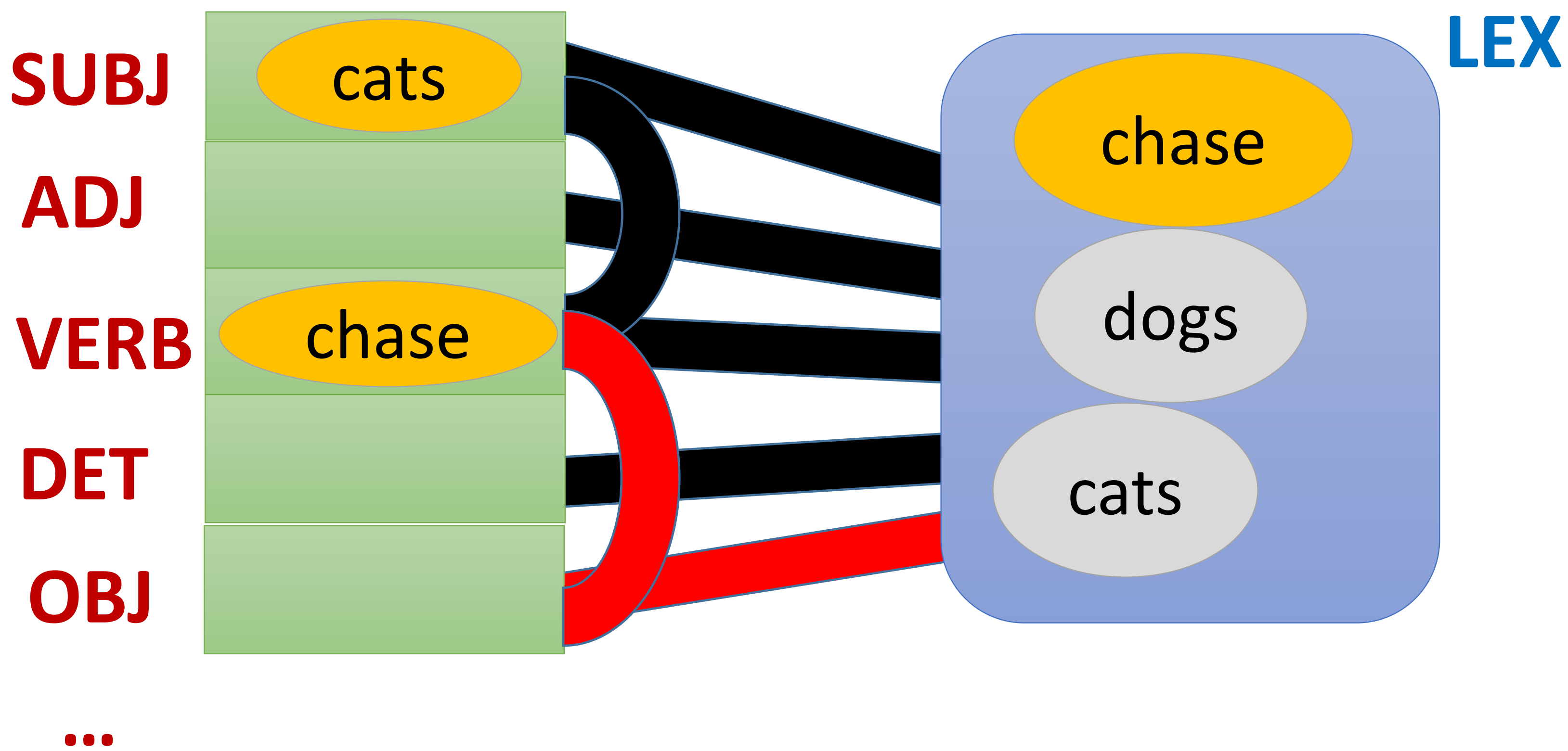
On input cats chase dogs



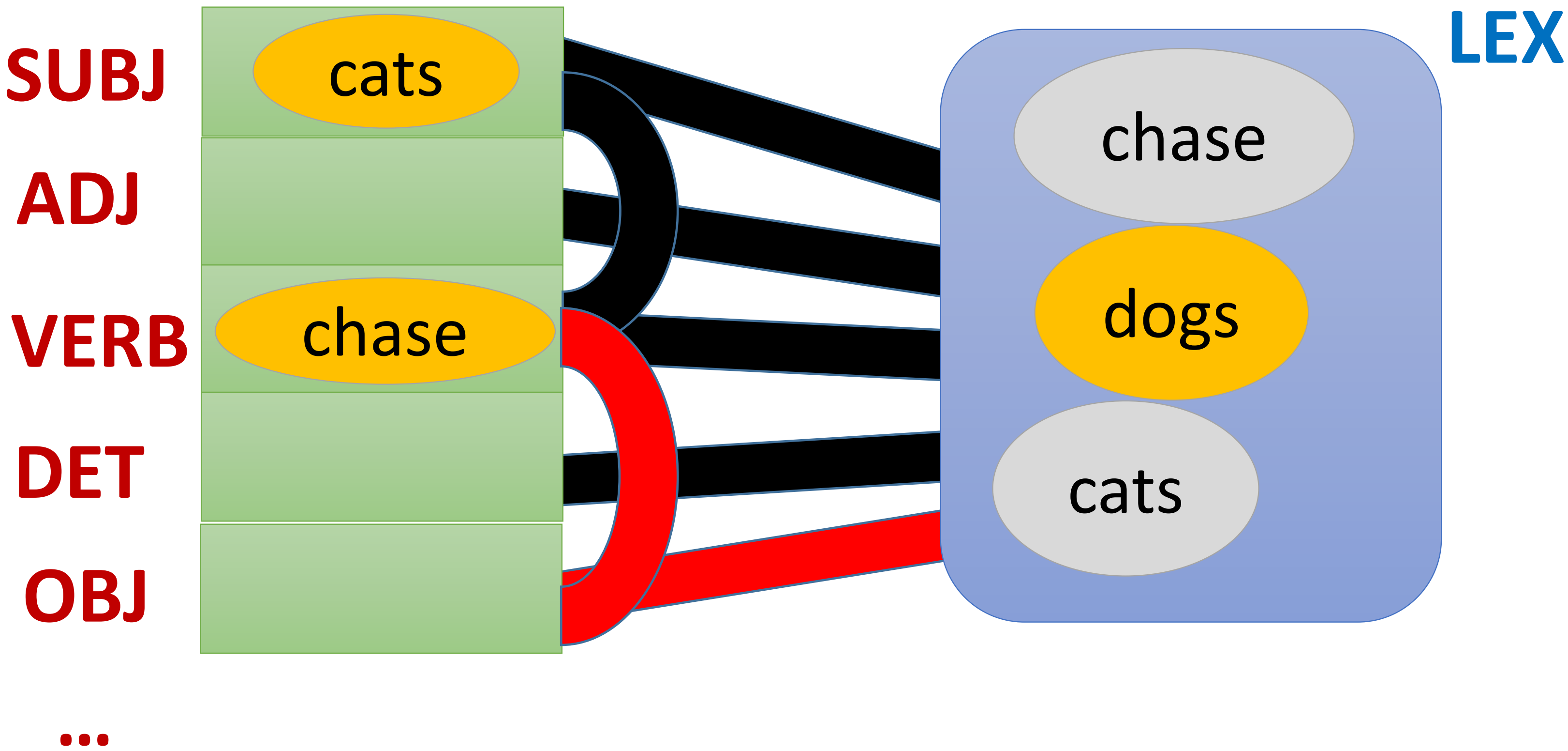
On input cats chase dogs



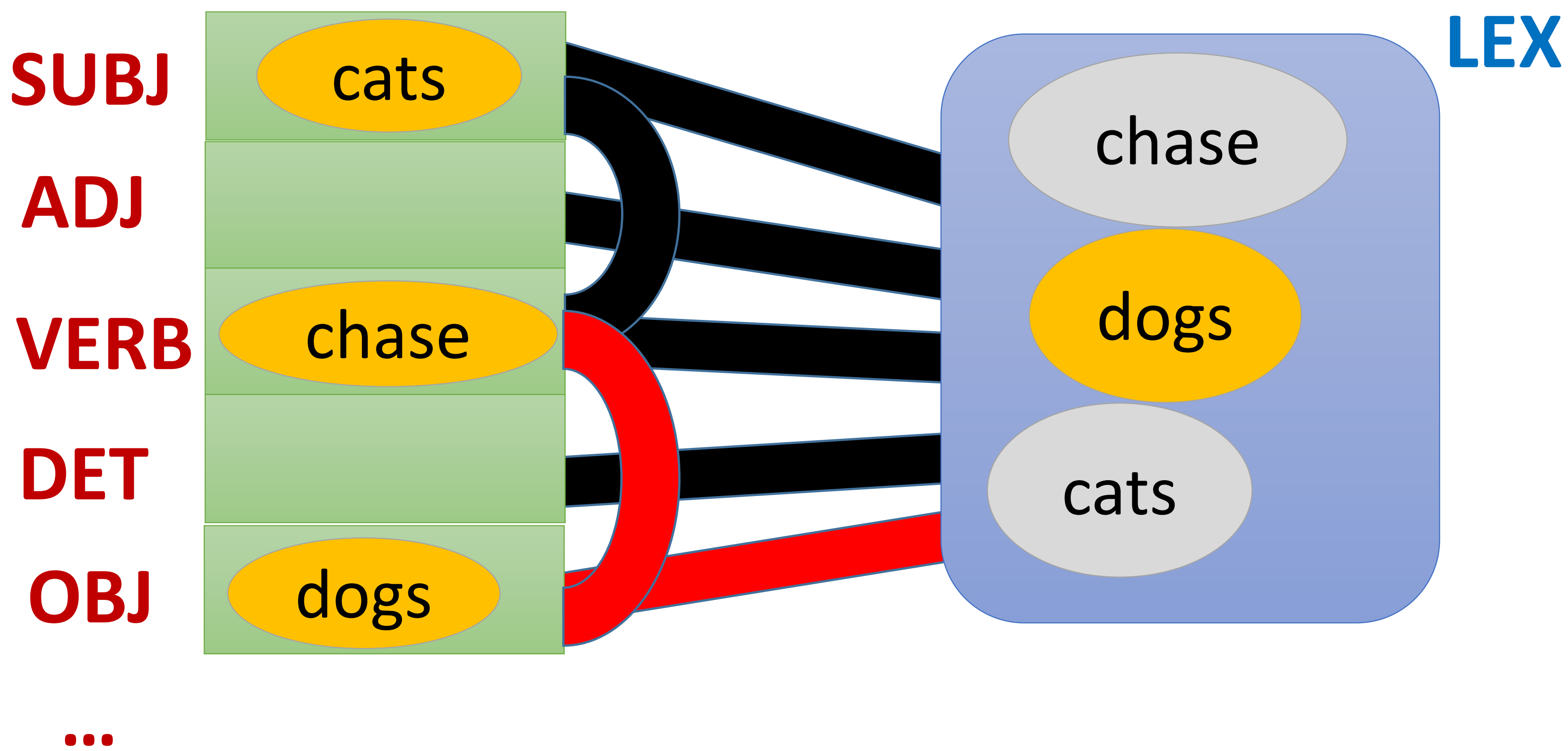
On input cats chase dogs



On input cats chase dogs

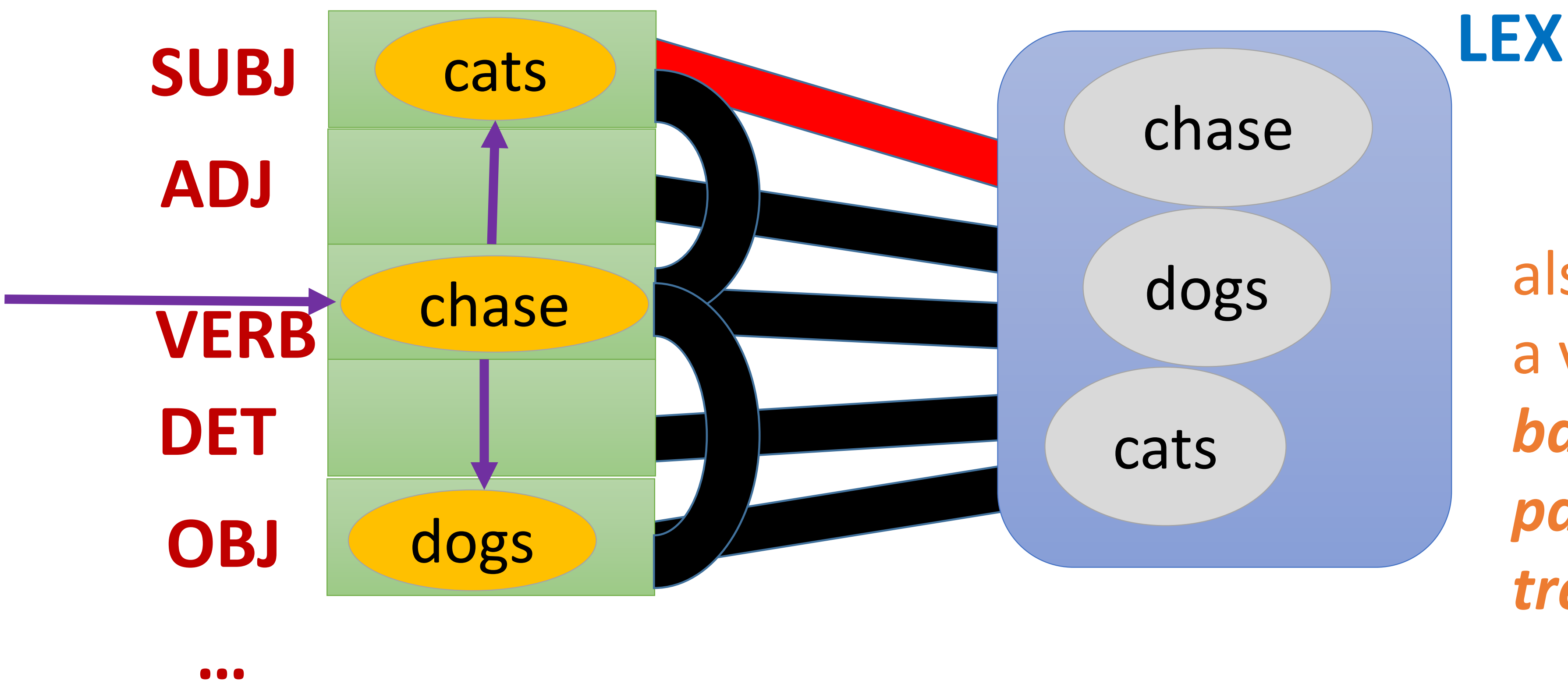


On input cats chase dogs



Q: But what does it mean for this device to correctly ***parse*** a sentence?

After input "**cats chase dogs .**" high synaptic weights form a *valid dependency tree*



also: a valid ***basic parse tree***

The Parser has limitations…

"the young couple who live in the next house saw the old little white car of the main suspect quite clearly"

???

(A parenthesis: Center embedding)

"cats, when they are fearless, chase dogs"

- *What is needed for center recursion?*
- Idea: parse on and return
- Note that, so far, the Parser was a ***finite state machine***
- State = the set of disinhibited fibers
- Fallback automaton: at any point, it can ***return*** to the most recent ***marked symbol,*** unmark it, then skip the ***already seen*** part, and continue parsing (NB: relies on sequence recall)

Theorem: FBA = CFL [NALOMA workshop 2022]