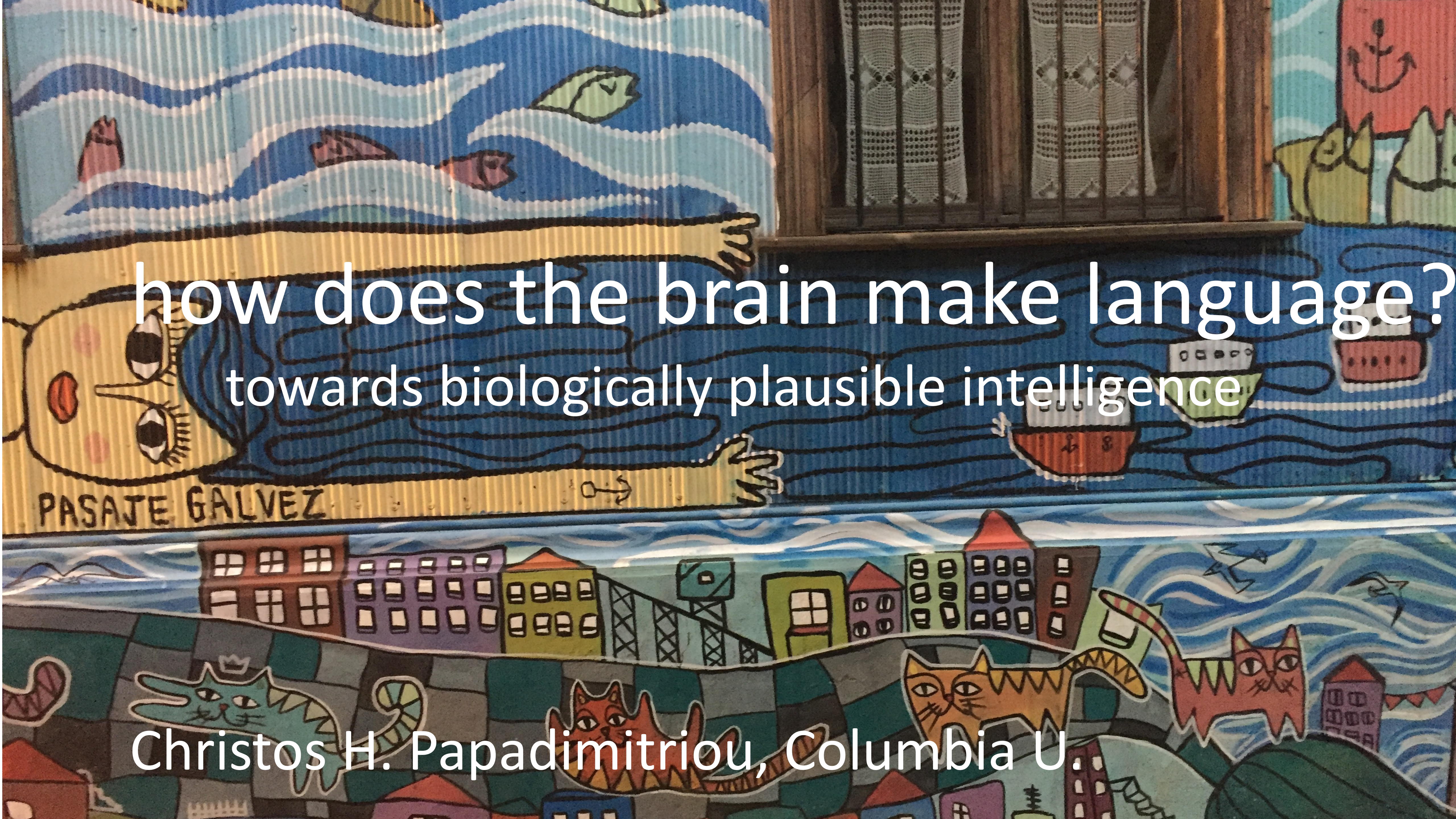
how does the brain make language?

towards biologically plausible intelligence



Christos H. Papadimitriou, Columbia U.

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

- Natural language interpretation and communication
- Common sense reasoning and understanding of the world
- Creative problem-solving and idea generation
- Flexibility and adaptability to new situations throughout a lifetime
- Emotional intelligence and understanding of social dynamics
- Ability to process information using wetware rather than hardware

[chat GPT, 2023]

But where does one start?

- ***Natural language interpretation and communication***
- Common sense reasoning and understanding of the world
- Creative problem-solving and idea generation
- Flexibility and adaptability to new situations throughout a lifetime
- Emotional intelligence and understanding of social dynamics
- Ability to process information using wetware rather than hardware

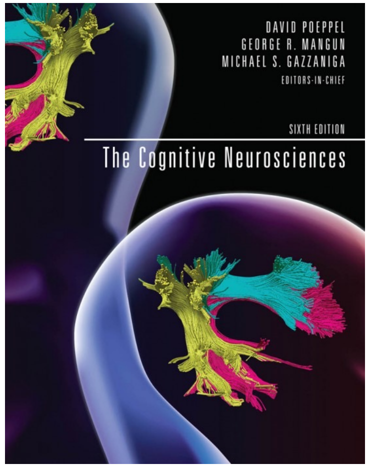
But where does one start?

- Natural language interpretation and communication
- Common sense reasoning and understanding of the world
- Creative problem-solving and idea generation
- ***Flexibility and adaptability to new situations throughout a lifetime***
- Emotional intelligence and understanding of social dynamics
- Ability to process information using wetware rather than hardware

Where does one start? My approach

- Natural language interpretation and communication
- Common sense reasoning and understanding of the world
- Creative problem-solving and idea generation
- Flexibility and adaptability to new situations throughout a lifetime
- Emotional intelligence and understanding of social dynamics
- ***Ability to process information using wetware rather than hardware***

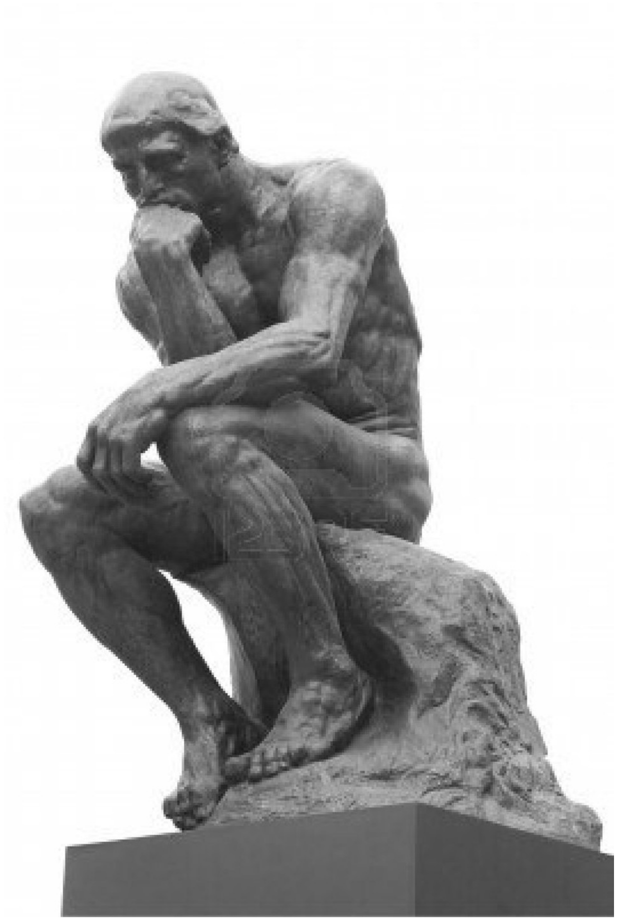
How does the Brain create the Mind?



*“…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

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



Our approach to Axel’s logic

Cognitive phenomena

↑ implements

math model of the brain

↓ **models reasonably well**

Spiking neurons and synapses

A math model of the brain

A finite number of brain areas

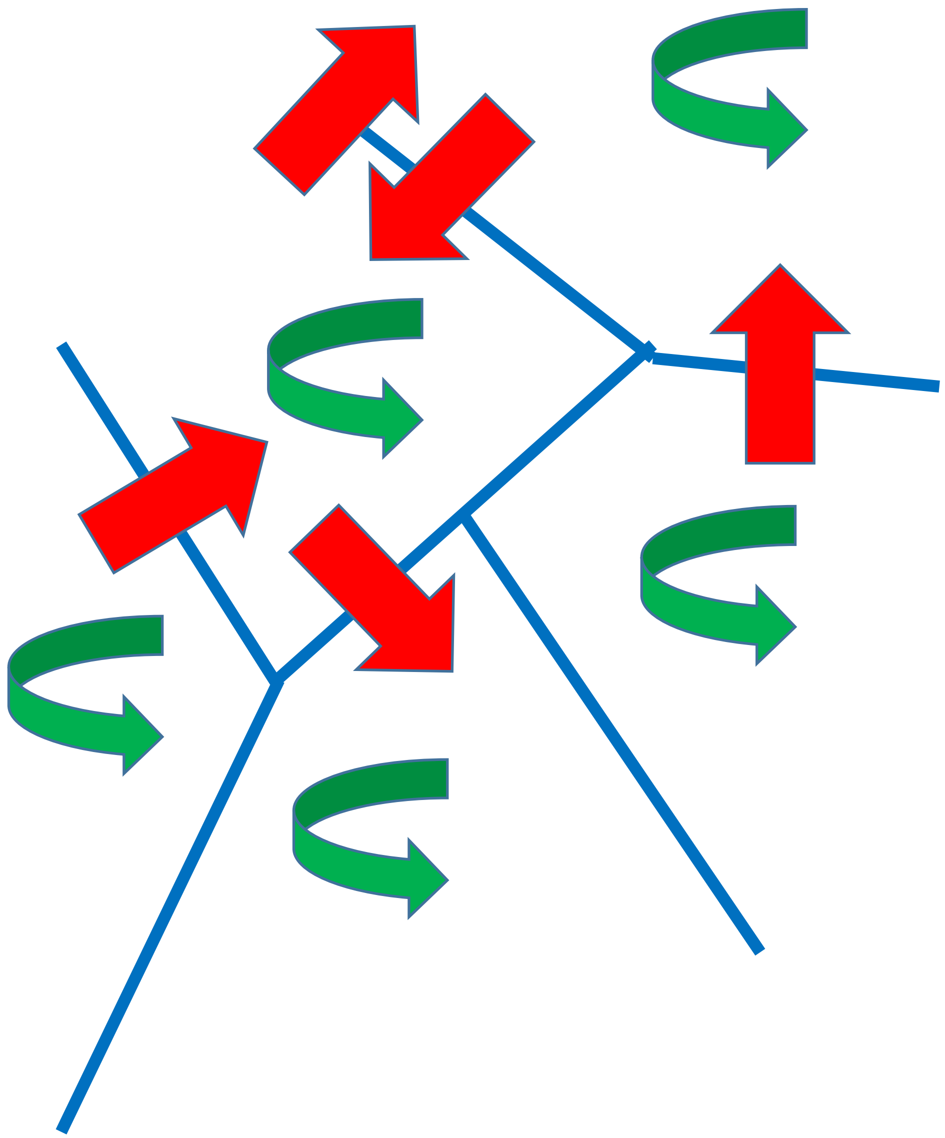
Each has ***n*** excitatory neurons

Certain pairs of areas are connected

by fibers of **random** connections

All areas are recurrently connected

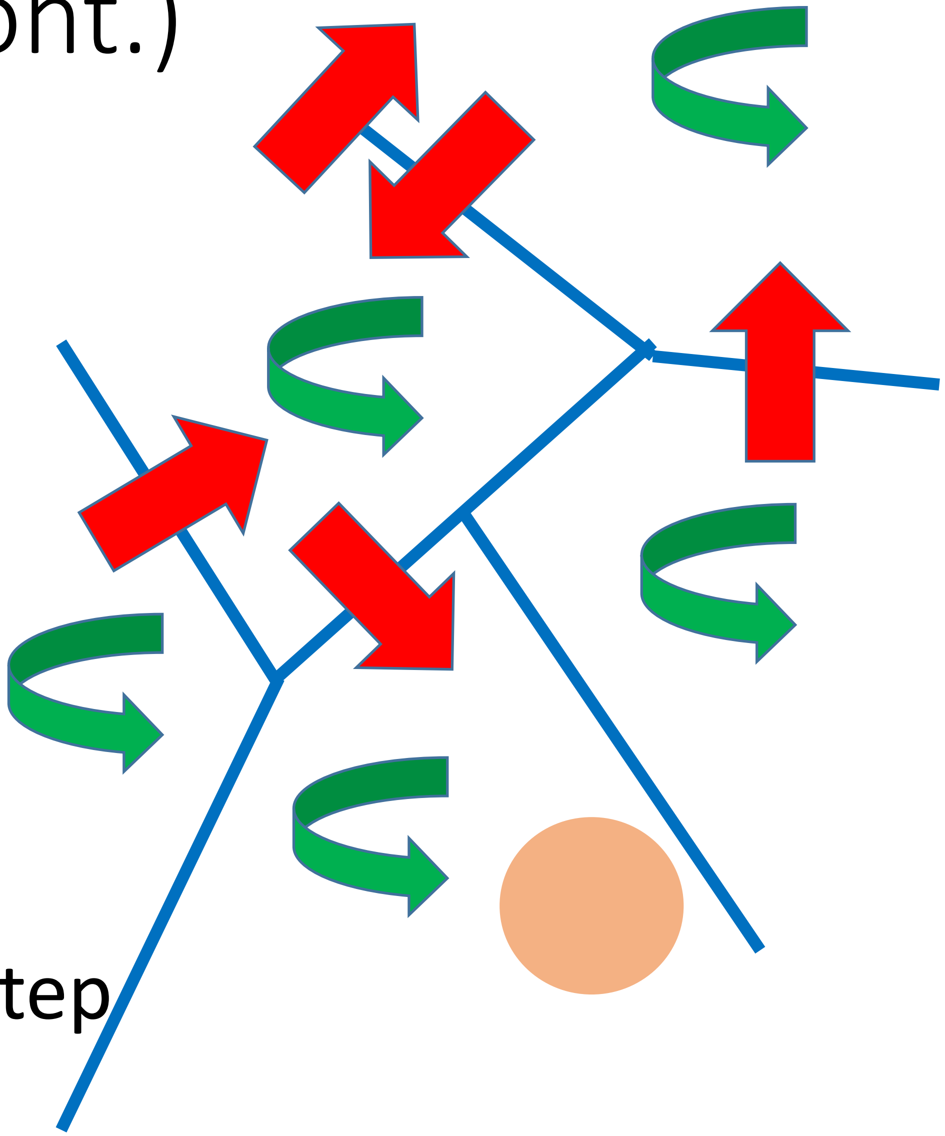
by **random** synapses



A math model of the brain (cont.)

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

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



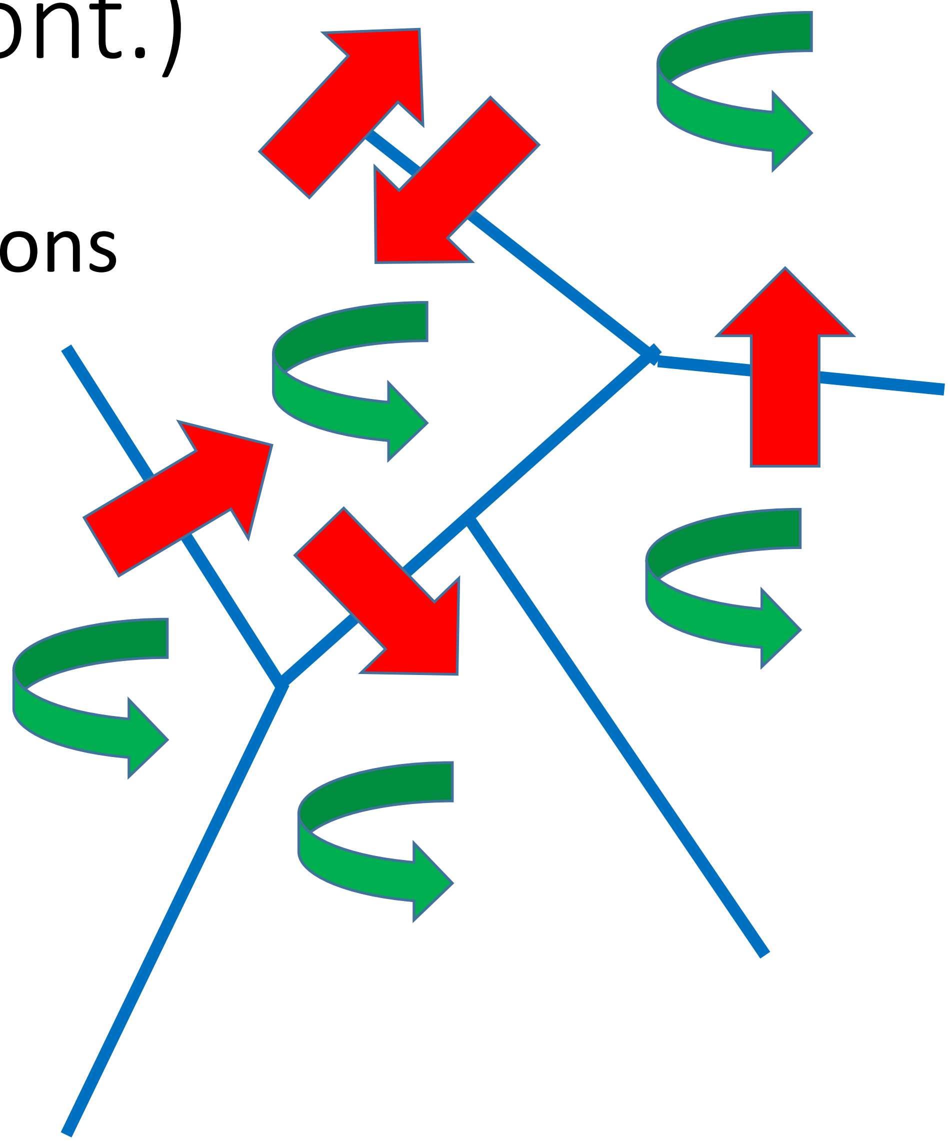
A math model of the brain (cont.)

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

Plasticity, Randomness, Selection

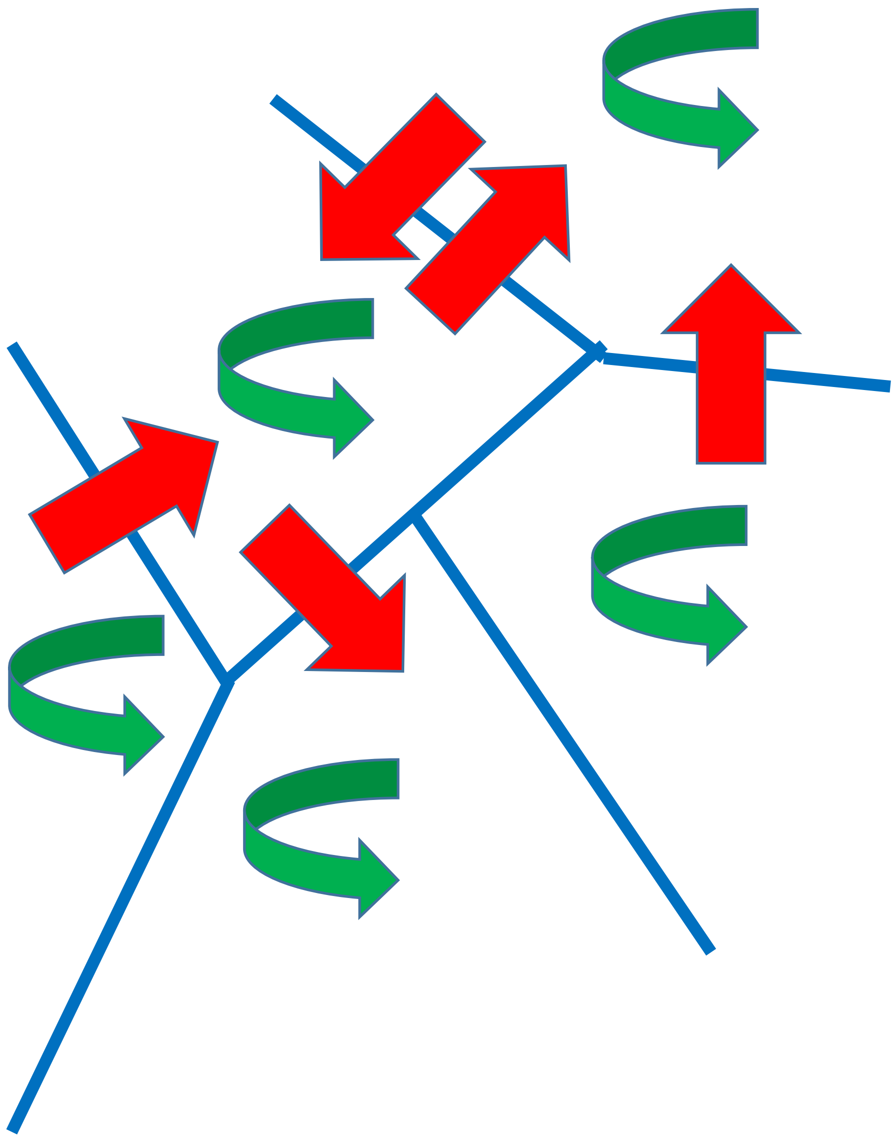
Typical Parameters:

$n \sim 10^7$, $k \sim 10^4$, $p \sim 10^{-2}$, $\beta \sim 5\%$



A math model of the brain

(Also: homeostasis, forgetting)
This defines a **dynamical system**
State: neurons that spiked, synaptic weights, inhibited areas
The next step function: fully defined!
Emergent behavior: ***assemblies of neurons***
Simulator available online
Q: But how is computation initiated?
A: Typically, by external stimuli



Short history of assemblies of neurons

- [Hebb 1949, Harris 2003, 2005; Buzsaki 2008, 2010, Yuste 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 operations

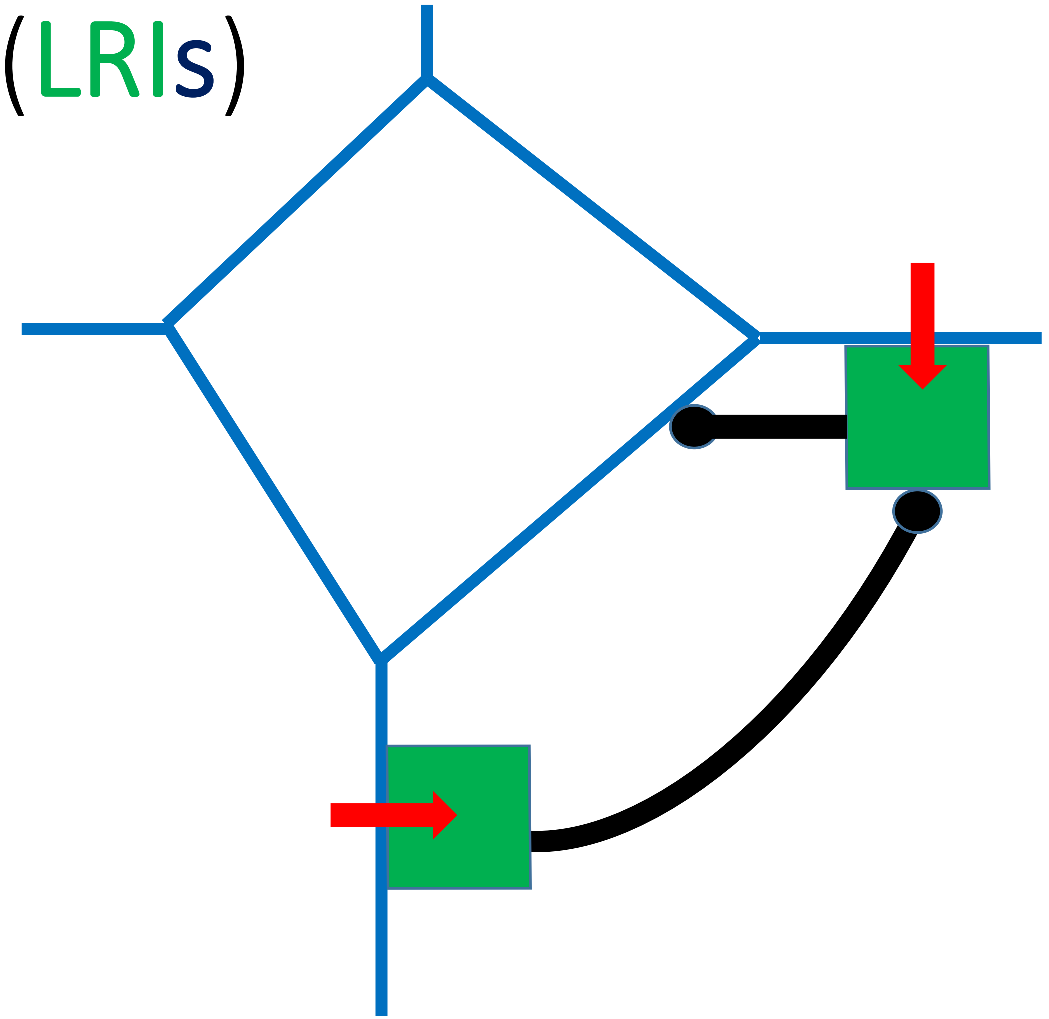
- Projection
- Reciprocal projection
- Association
- Pattern completion
- Merge
- Sequence recall
- Few shot learning of simple classification tasks

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

*Plus, **simulations***

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

- Populations of inhibitory neurons with long axons
- 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 = computation

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

The math model 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 20
- A ***lazy simulation*** technique reduces this to **$O(pk^2T^2)$**
- 10^6 speedup, allows us to simulate ***a few seconds*** of brain time at 3 GHz
- Enough for some cool cognitive phenomena

Recall: Computation in the brain

- Cognitive phenomena

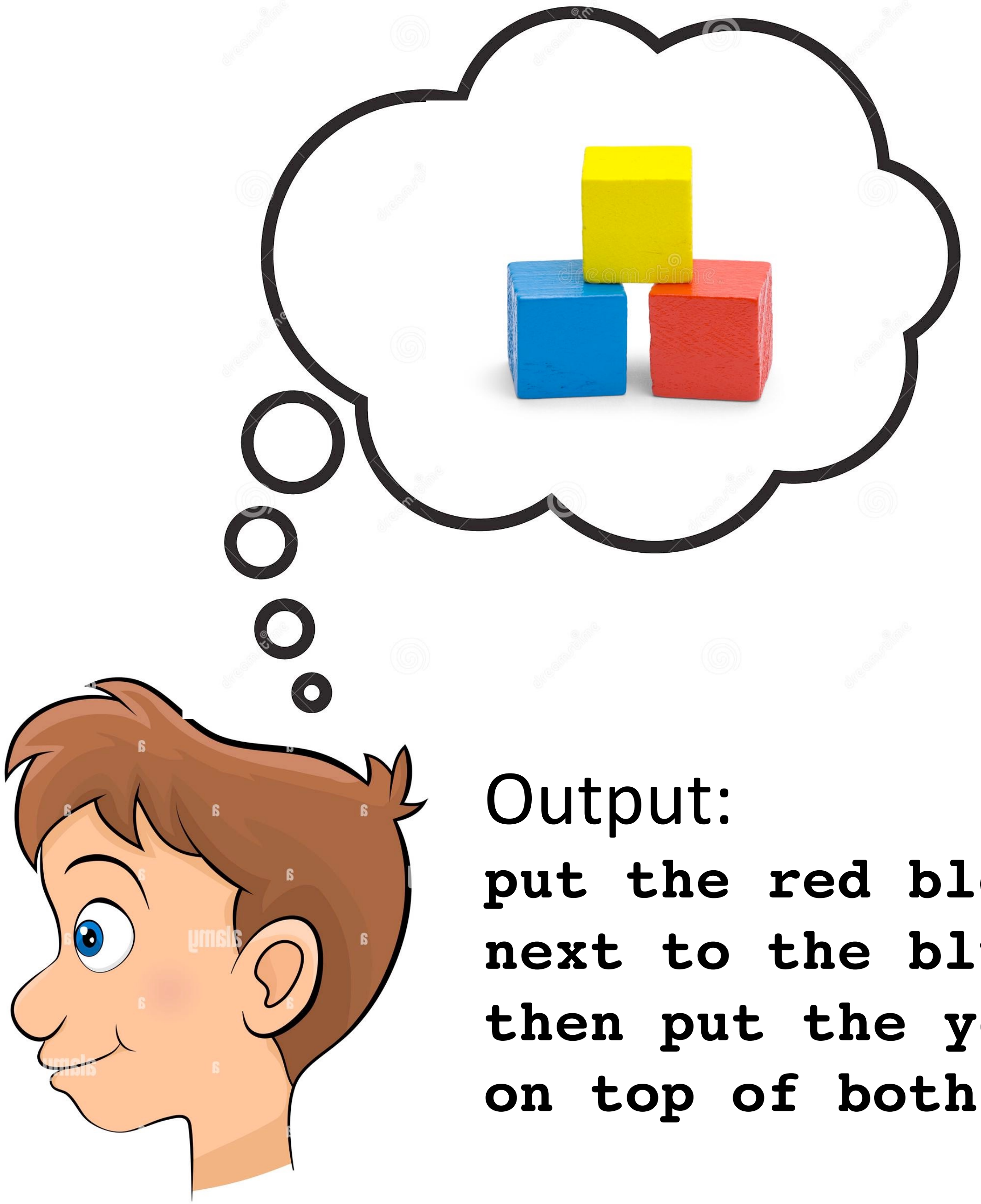
implements *next*

math model

models reasonably well

spiking neurons and synapses

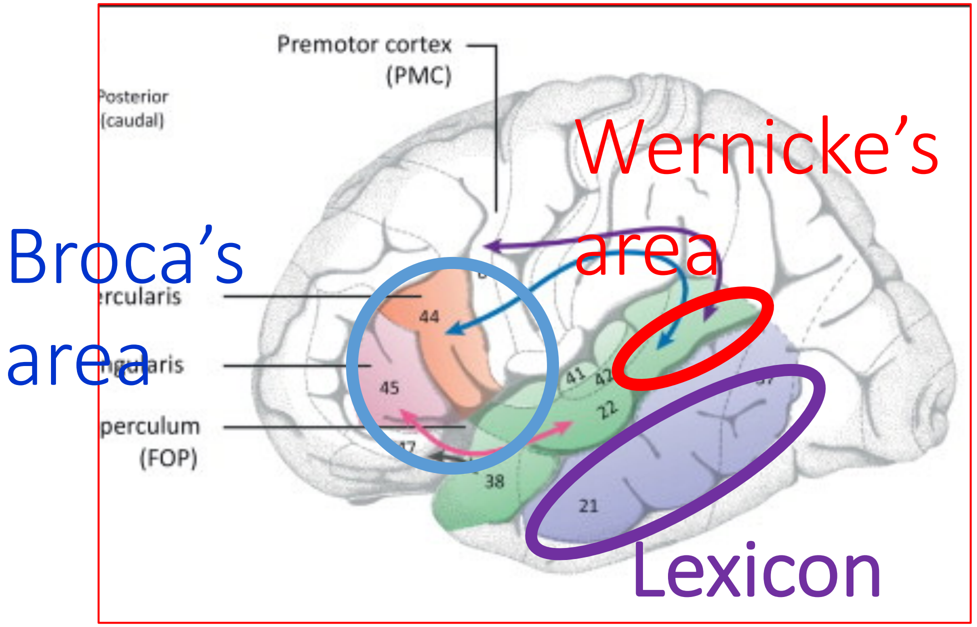
Planning in the Blocks World (AAAI 2021)



Output:

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

Language!

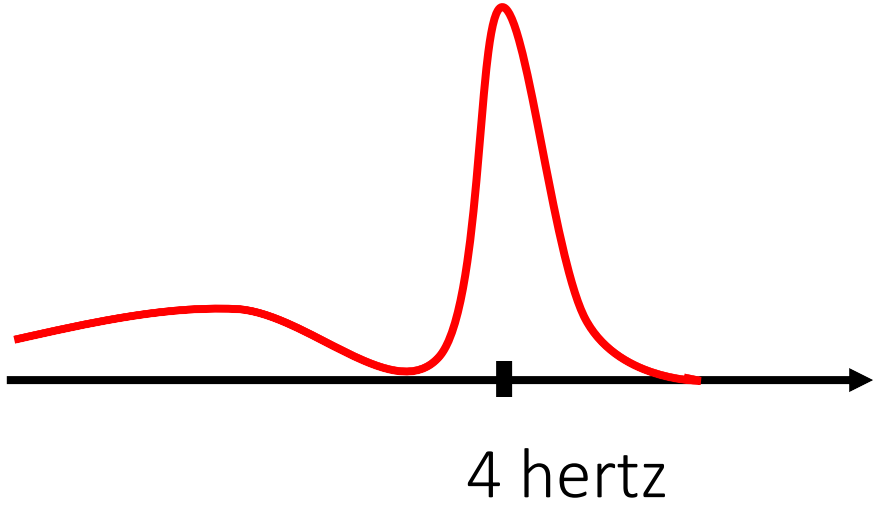


It is the hardest thing any brain has done, and so it must hold the key

But how does it happen in the Brain?

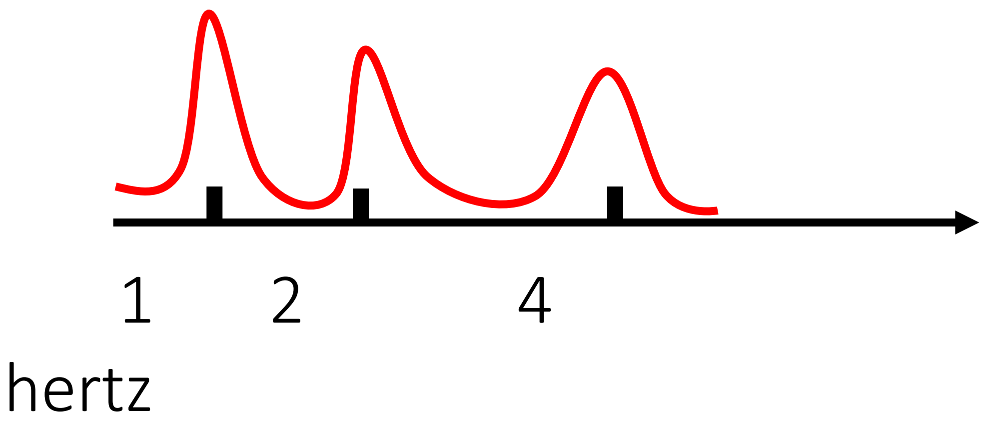
The [Poeppel 2016] experiment

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

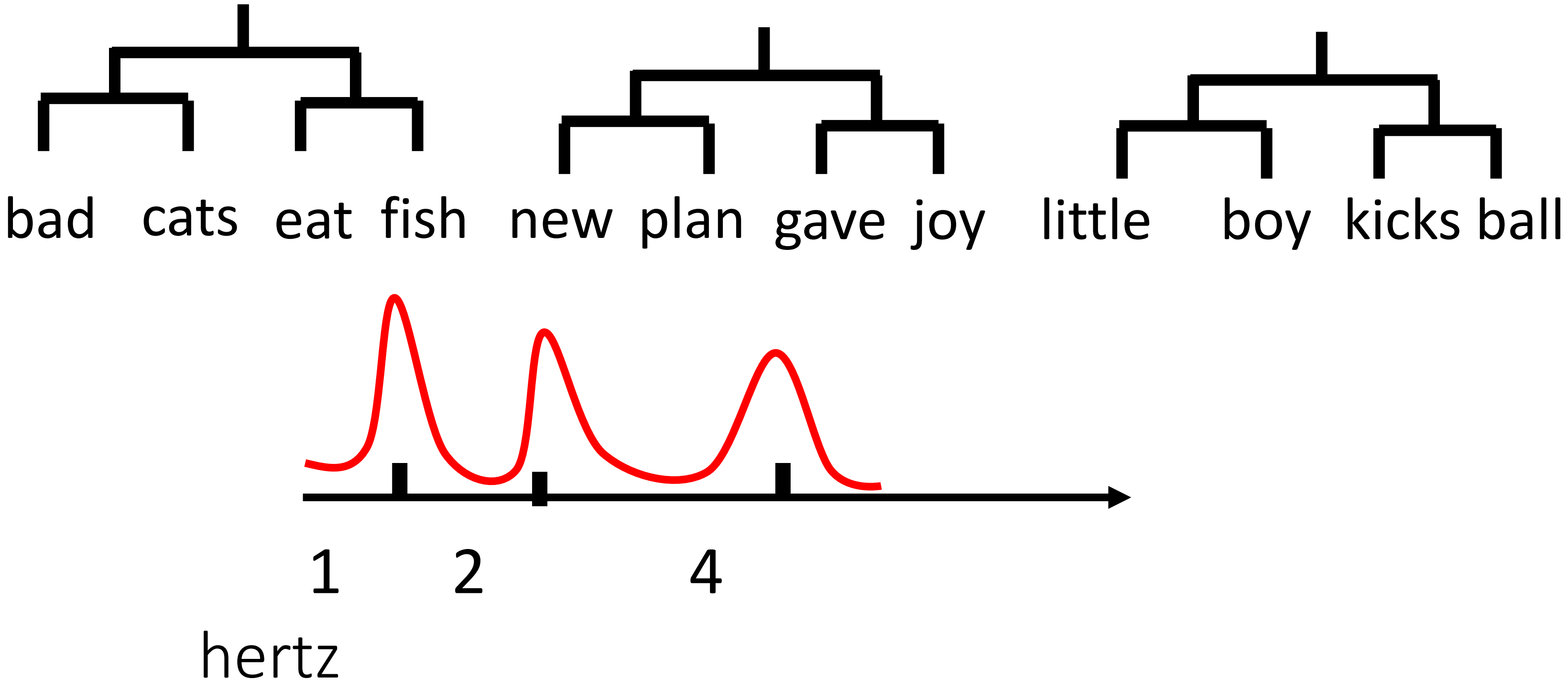


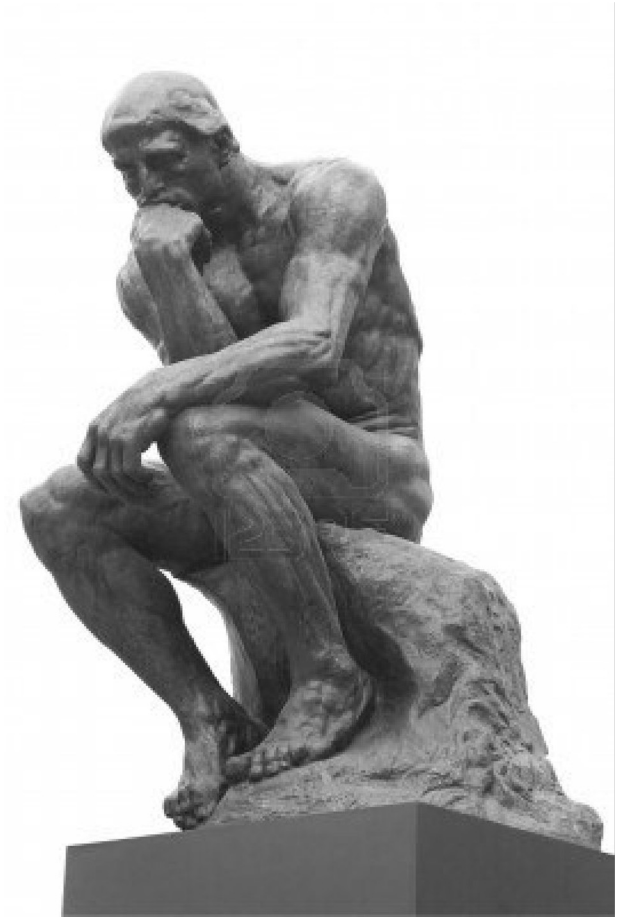
The [Poeppel 2016] experiment, stage II

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



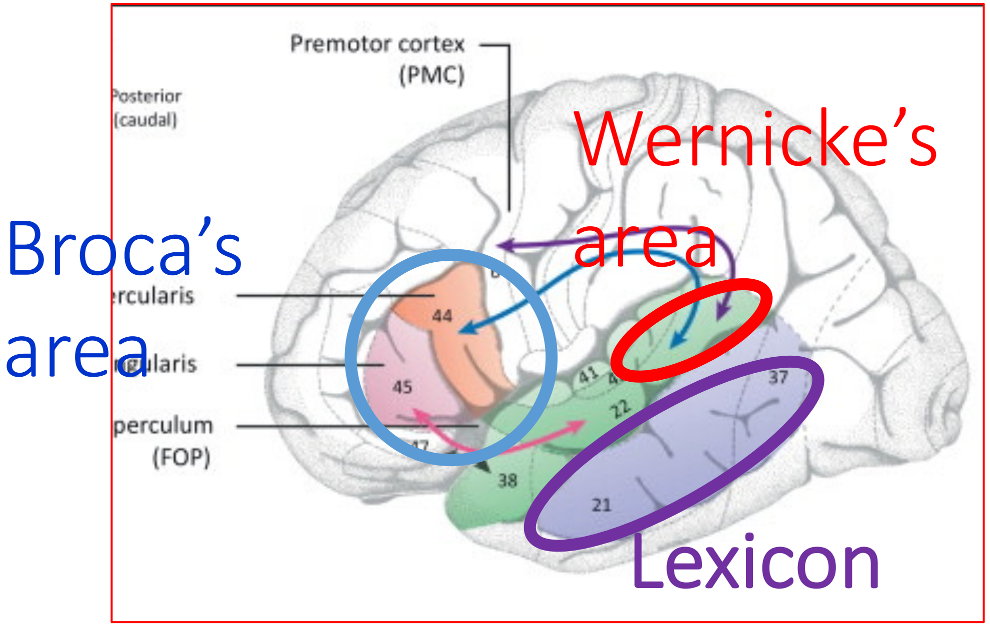
My interpretation





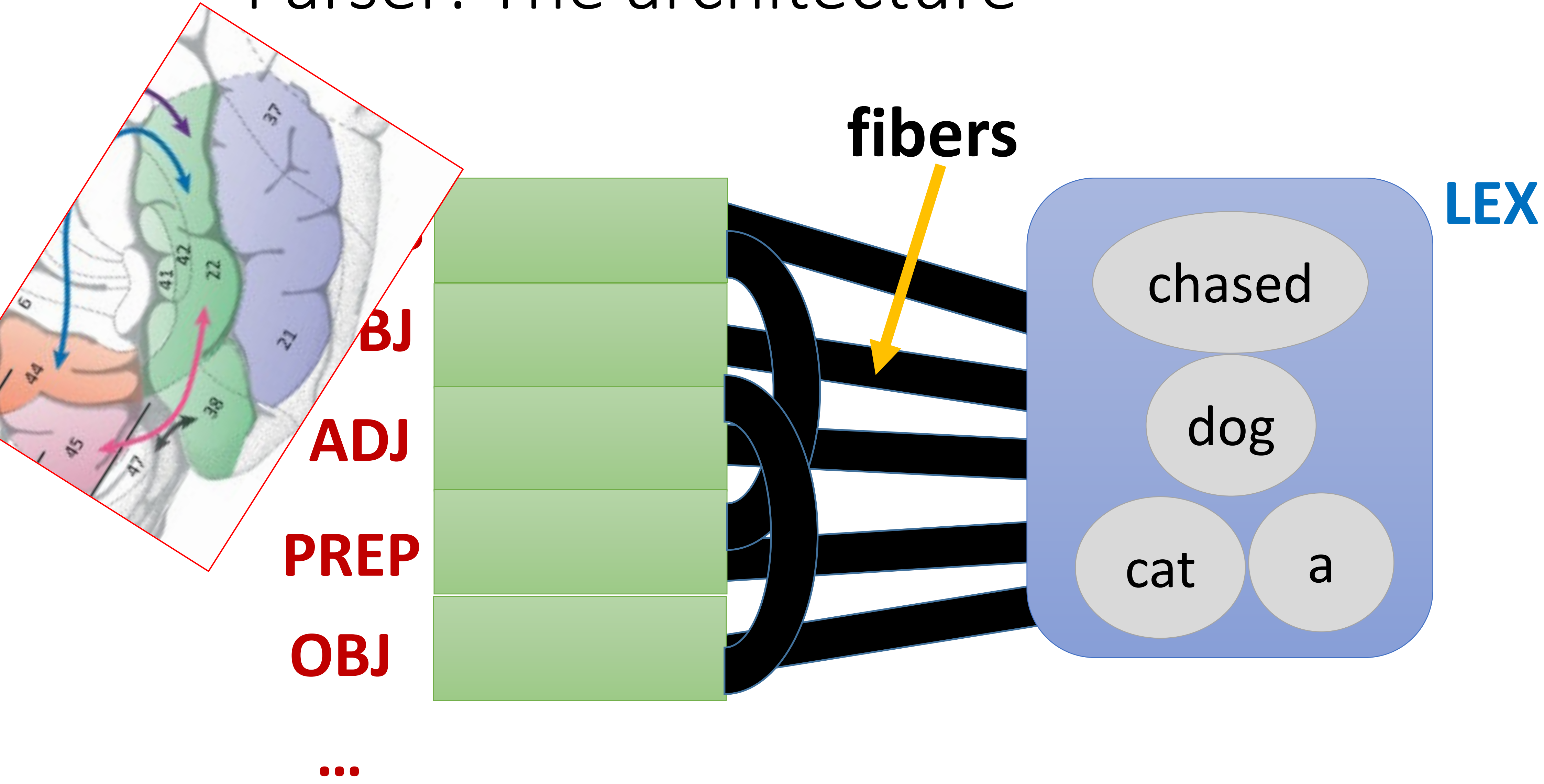
- Hmmm…. how are these tree representations created?
  in about a dozen spikes per step?
- (Why a dozen?
  neuron spikes ~ 50Hz :
  language ~ 4Hz)
- Can we simulate parsing?

Next:



a parser of English implemented by spiking neurons

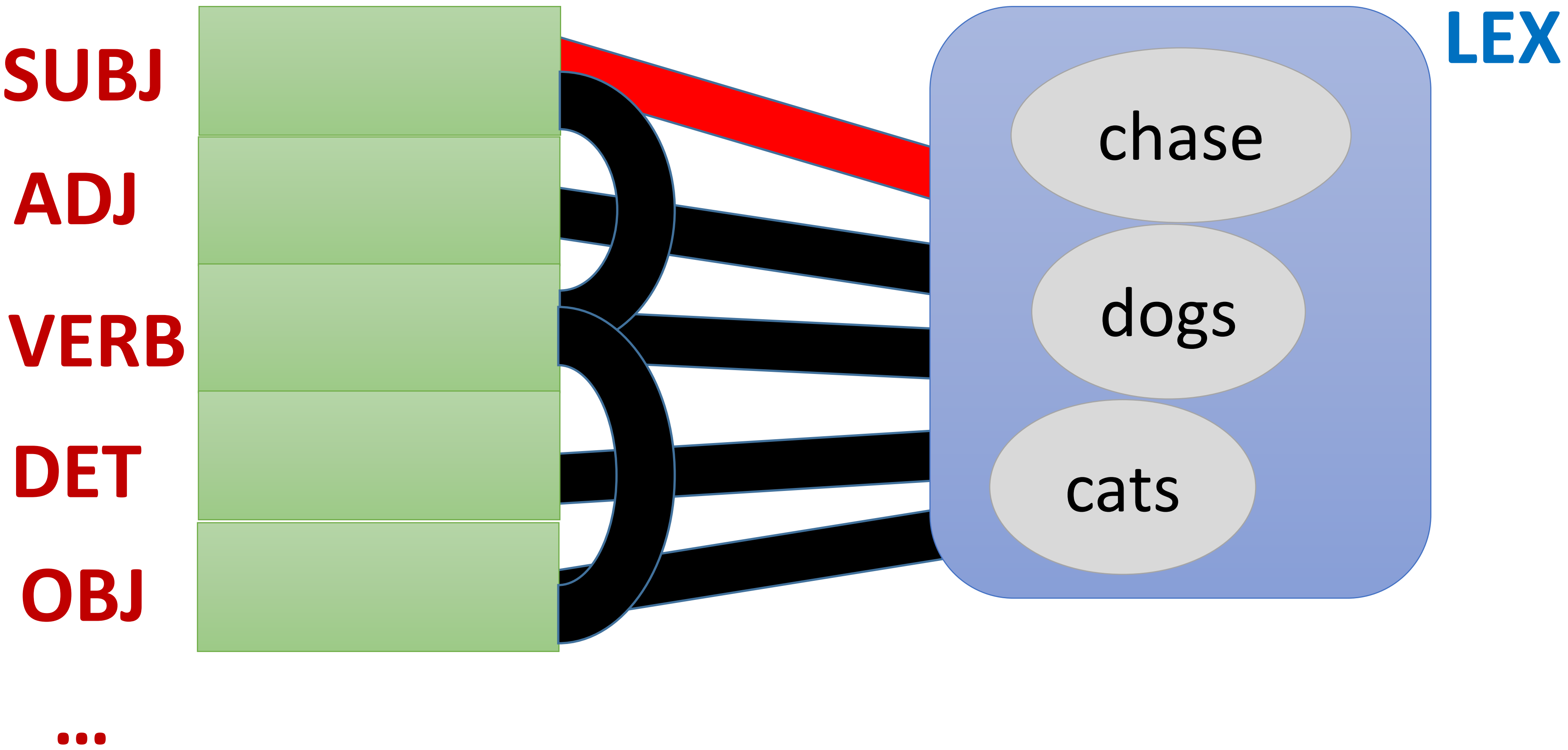
Parser: The architecture



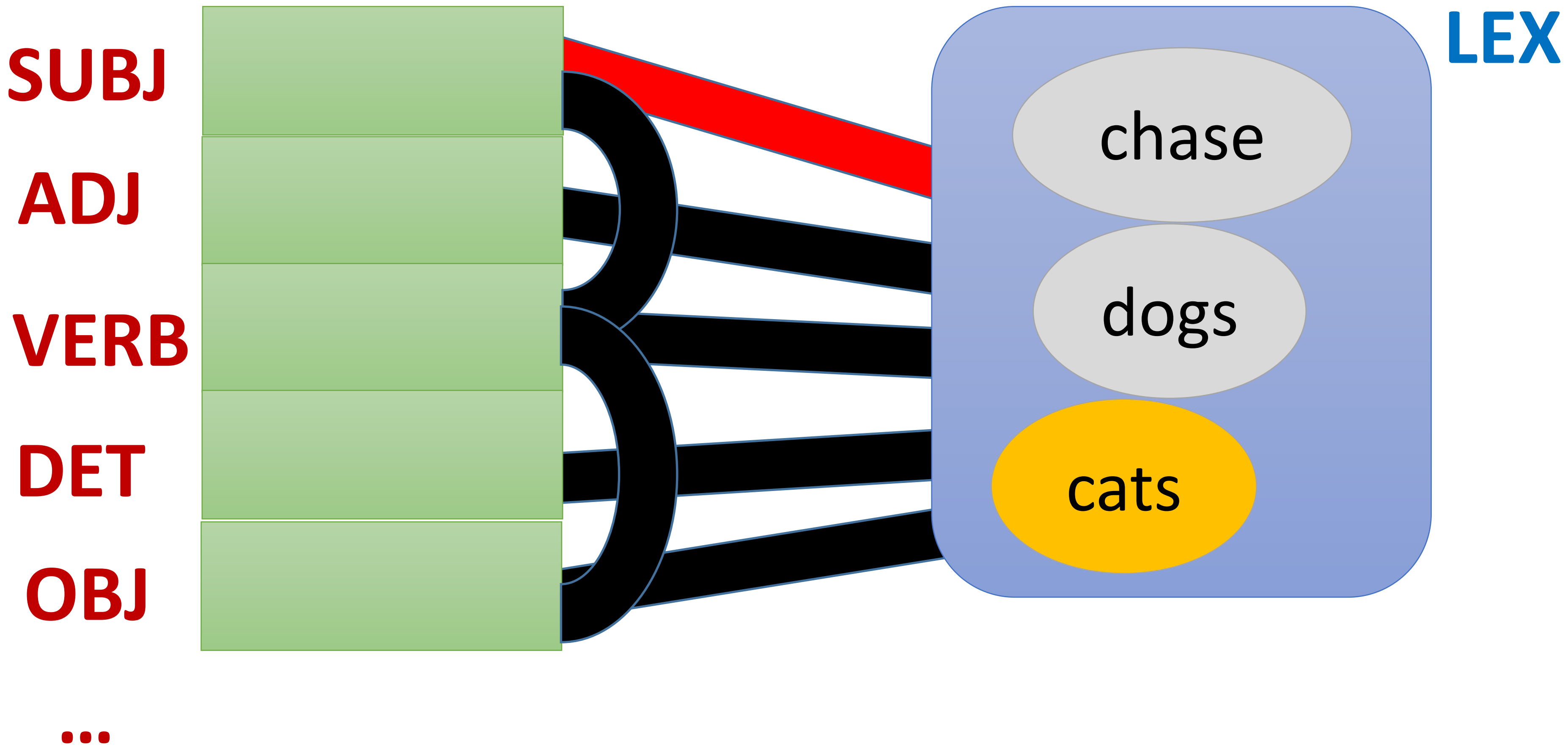
The Parser

- Input: a sequence of excitations of word assemblies
- (We assume that **phonetics** *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***

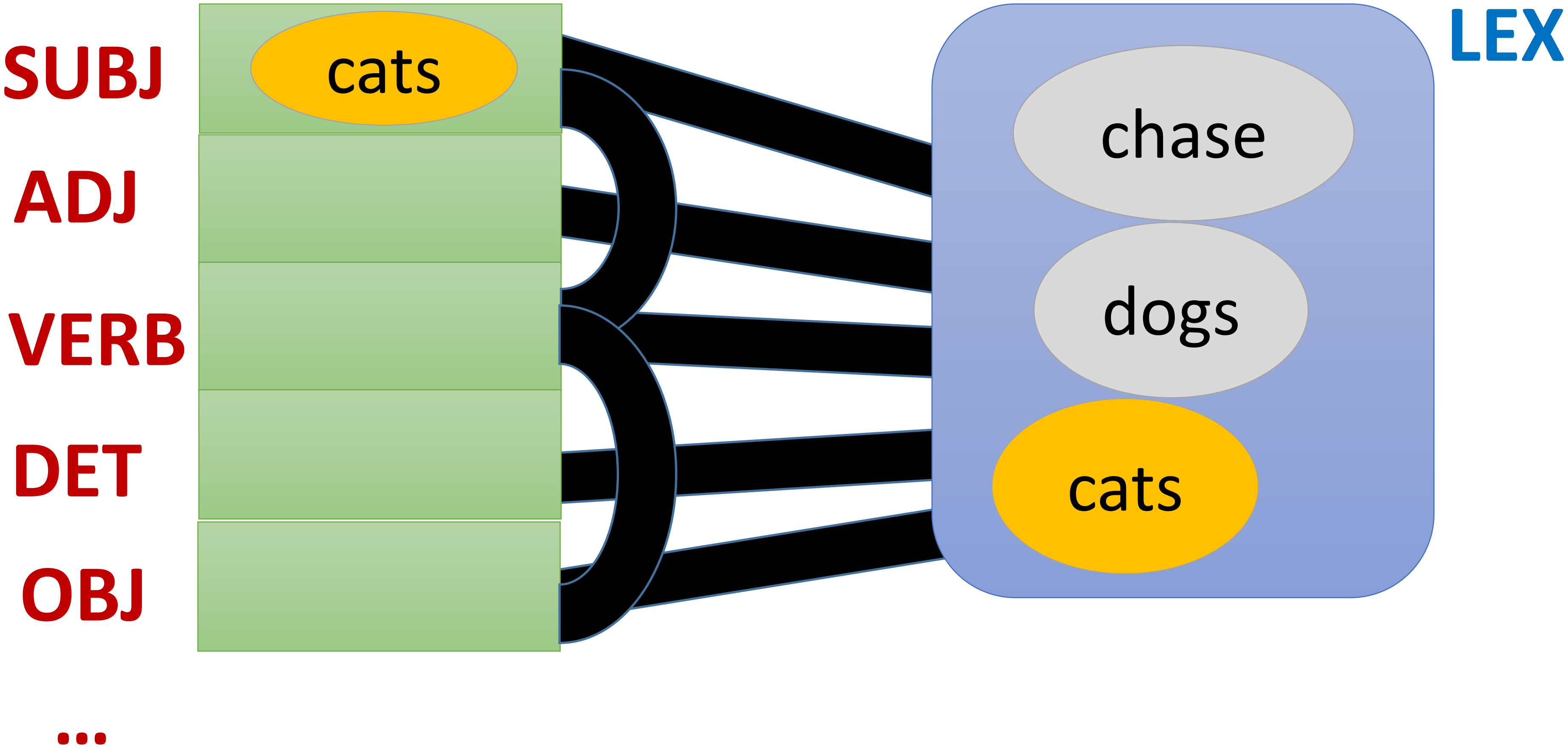
On input **cats chase dogs**



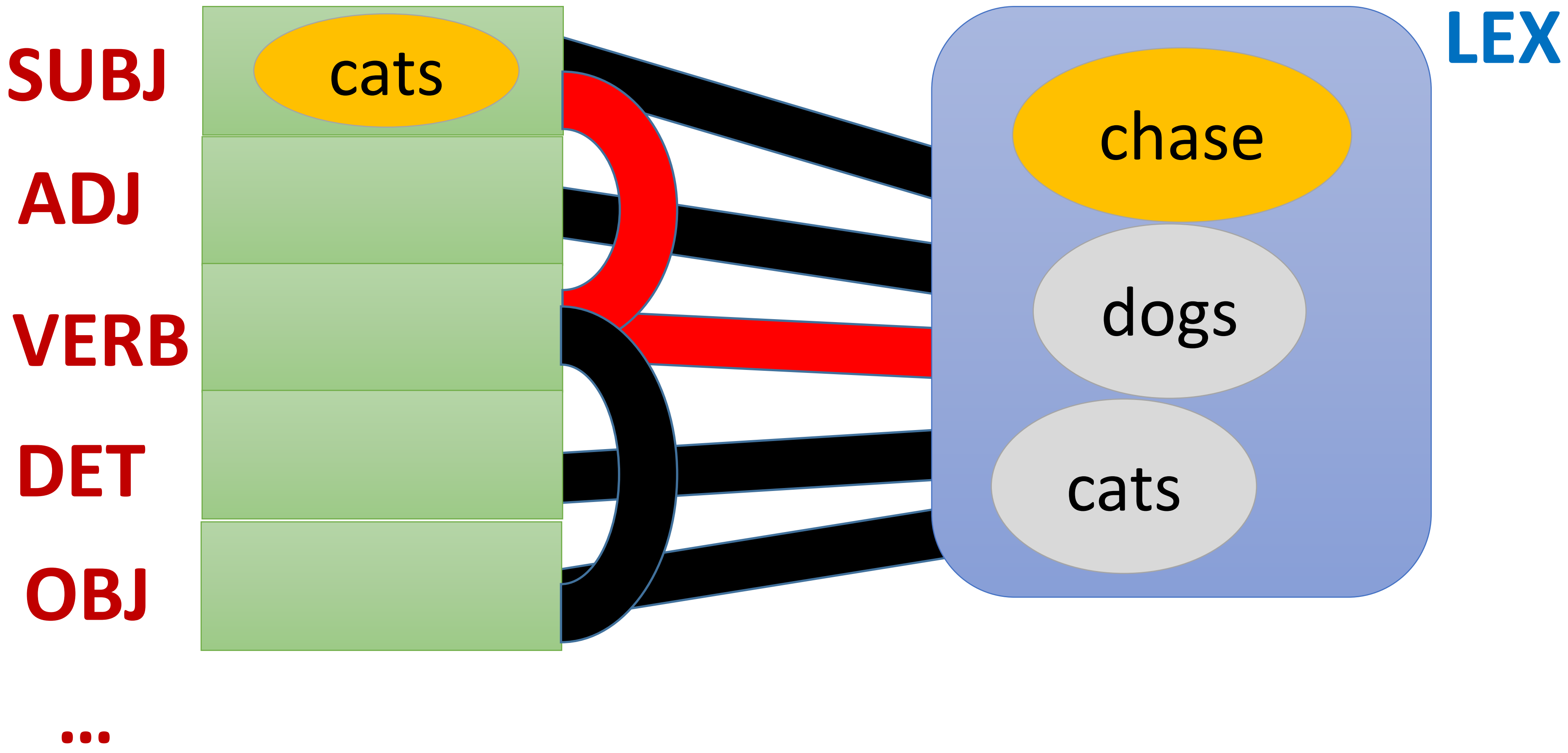
On input cats chase dogs



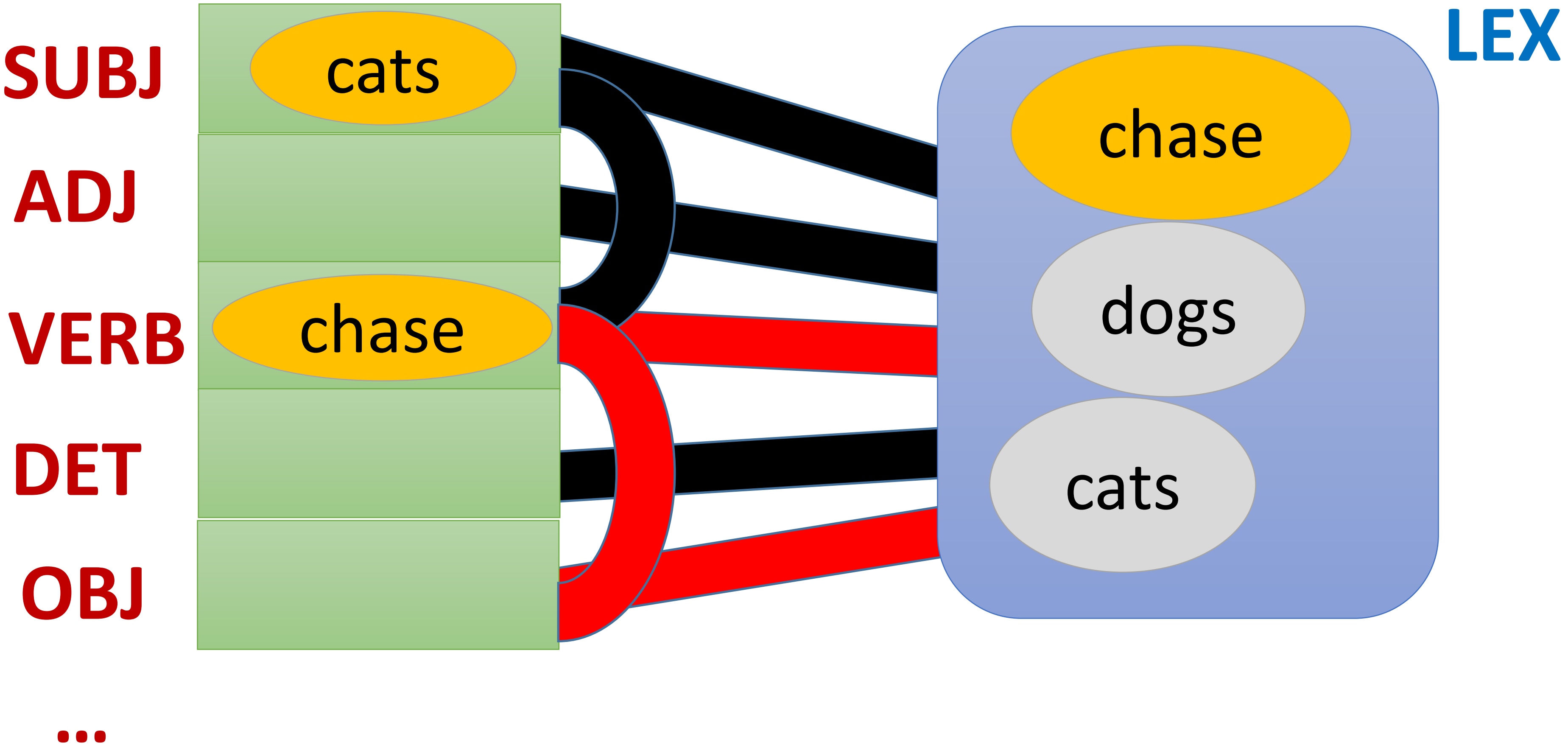
On input cats chase dogs



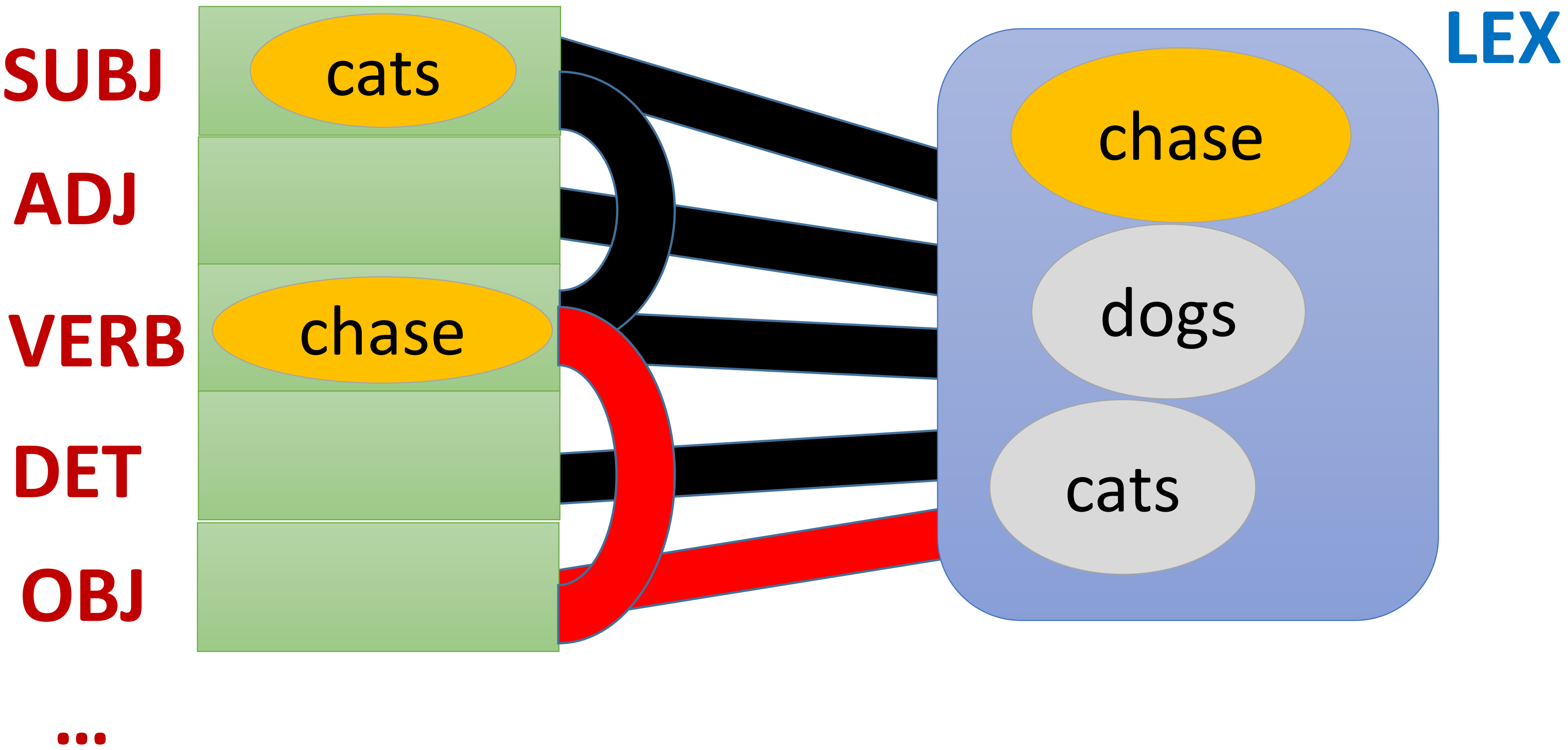
On input cats chase dogs



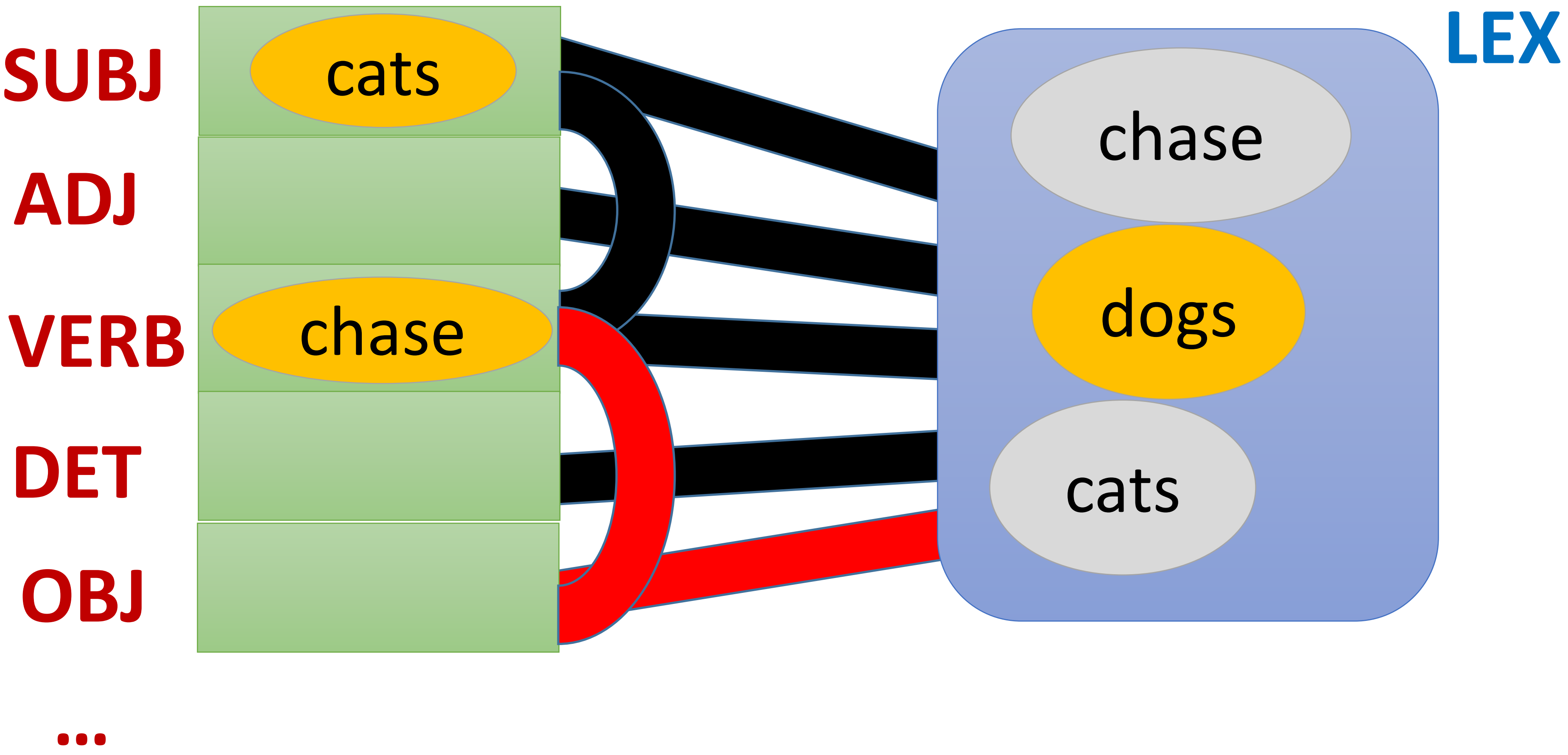
On input cats chase dogs



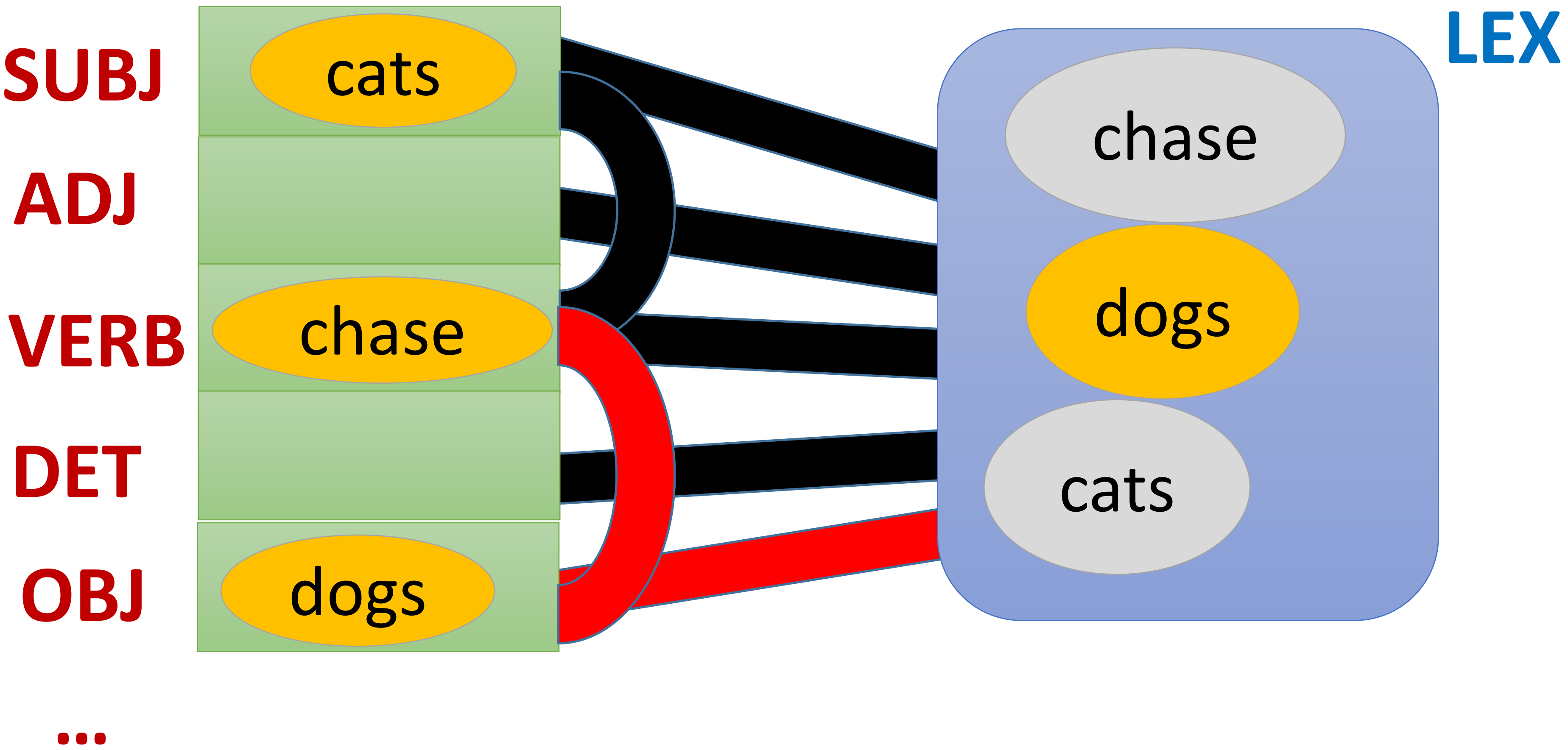
On input cats chase dogs



On input cats chase dogs

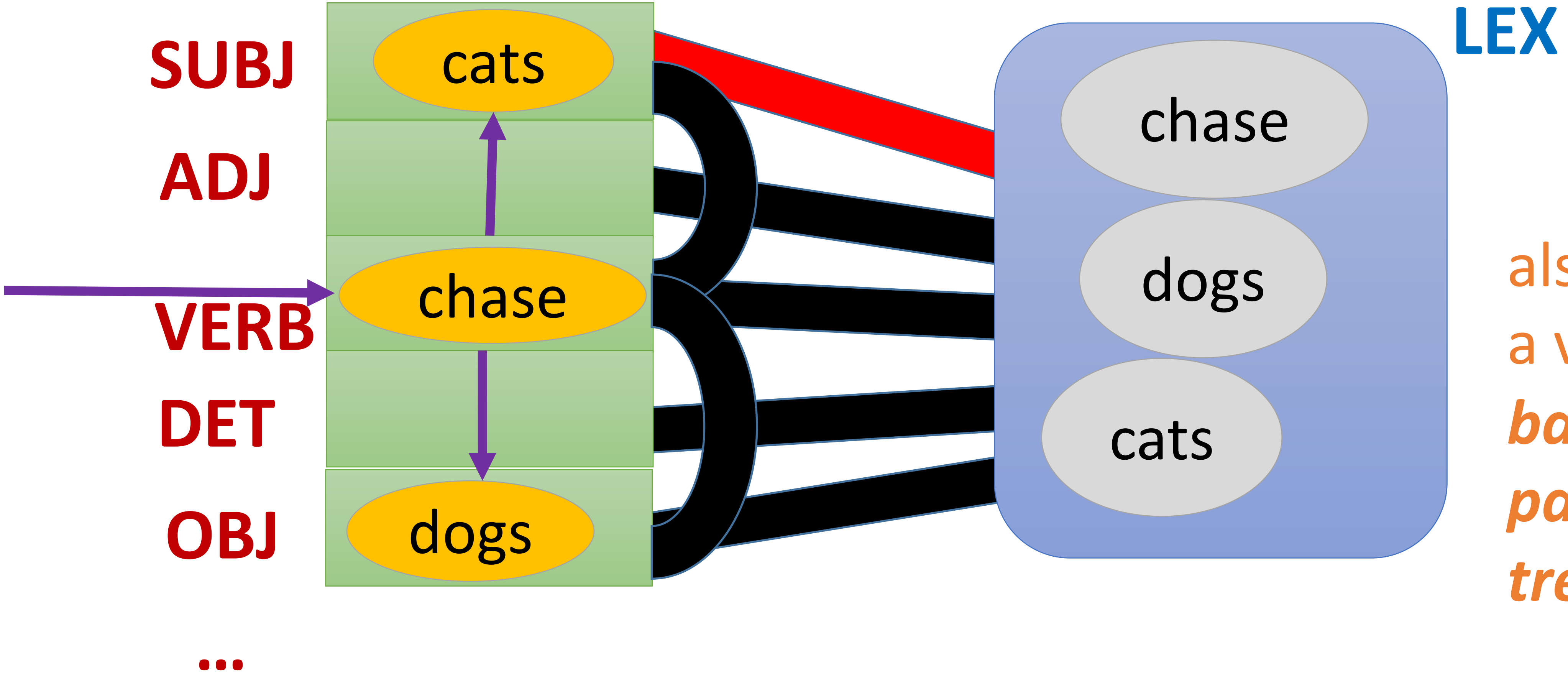


On input cats chase dogs



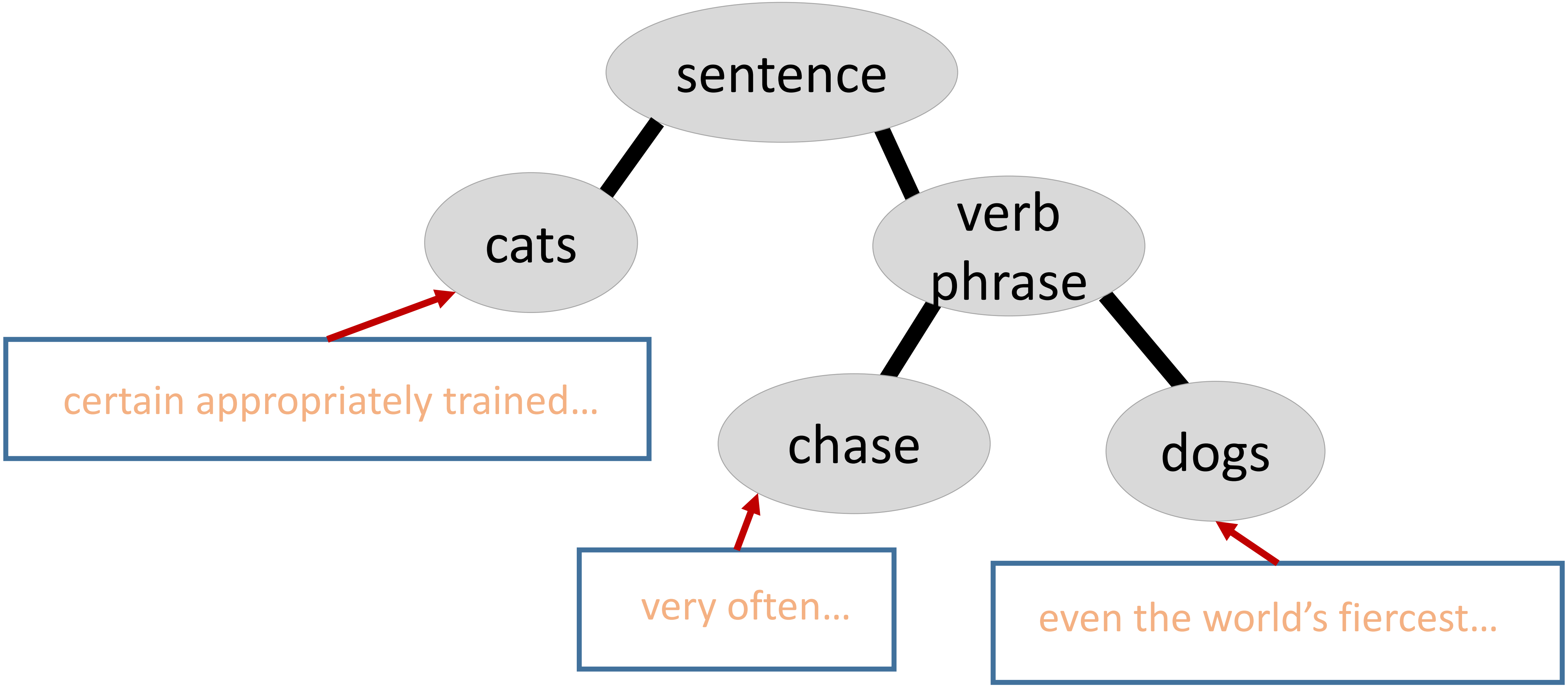
Q: But what does it mean for this device to correctly ***parse*** a sentence?

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



also:
a valid
basic parse tree

Basic parse tree



The Parser as neuromorphic computation

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

The Parser

- 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 running: simple parsers of ***Russian, Japanese, Hungarian***, ***Chinese***
- Paper in ACL 2022, code available online

The Parser

- It parses simple sentences such as:

“the young couple who live 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 running: simple parsers of ***Russian, Japanese, Hungarian***, ***Chinese***
- Paper in ACL 2022, code available online

(A parenthesis on center embedding)

“cats, when they are fearless, chase dogs”

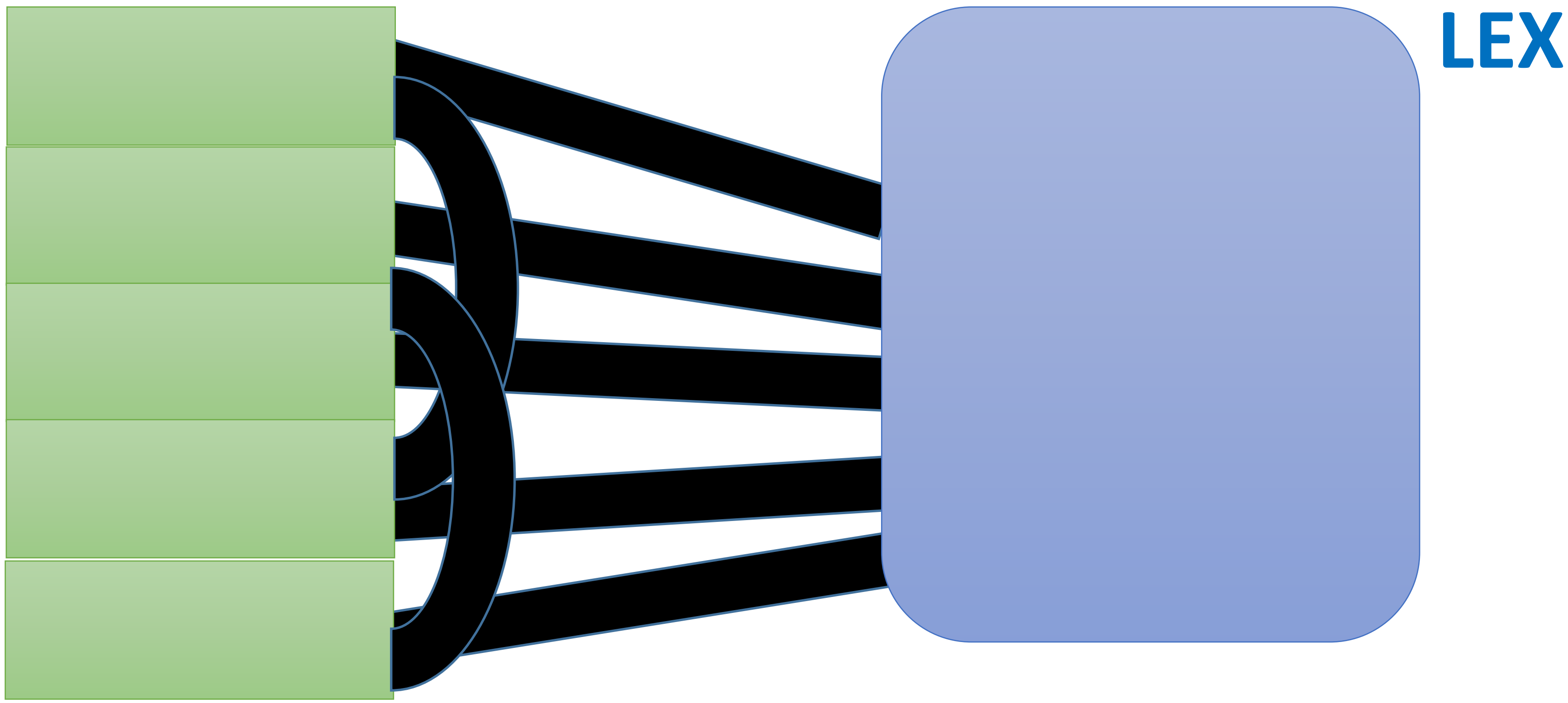
- *What is needed for center recursion?*
- Idea: parse on and return
- Note that, so far, the Parser was a ***finite state machine***
- State = the graph of disinhibited fibers
- Fallback automaton: at any point, it can ***return*** to the most recent ***marked symbol***, unmark it, then skip the ***already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

Theorem: FBA = CFL [NALOMA workshop 2022]

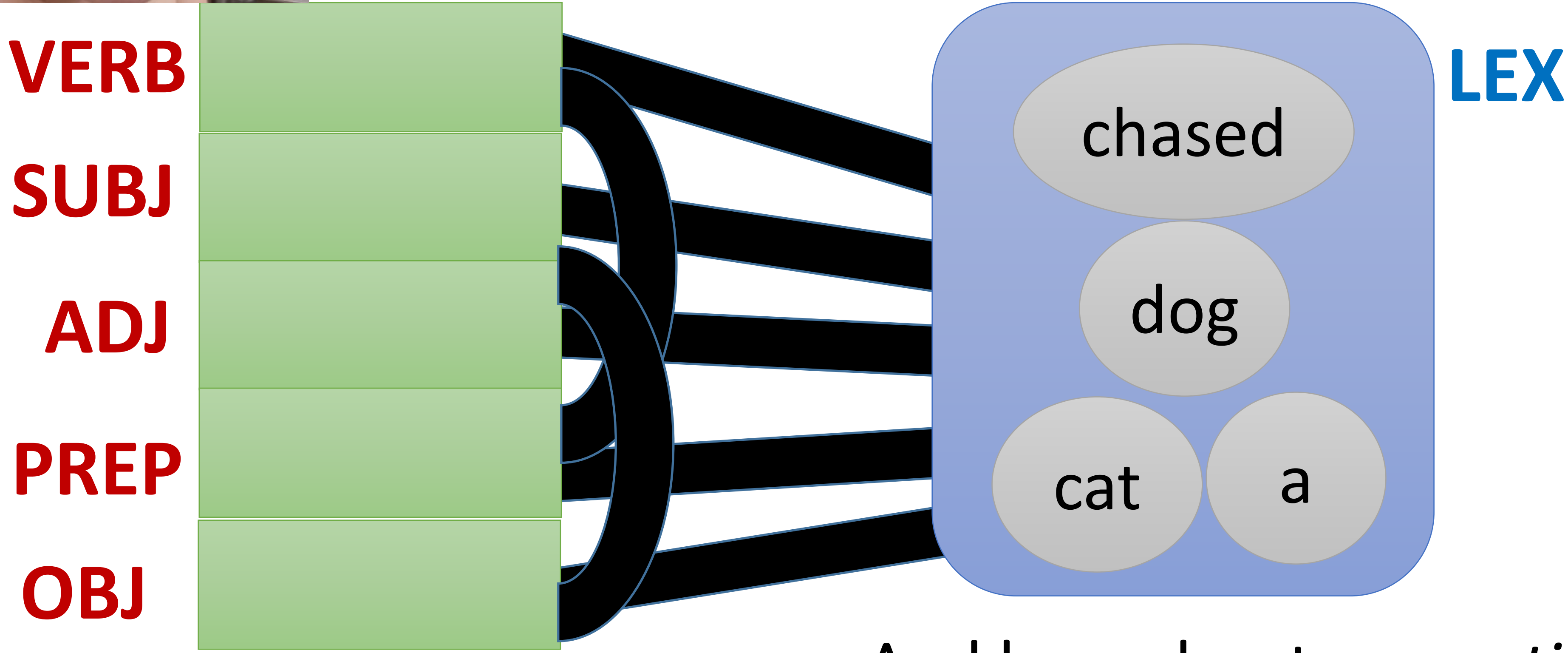
(On methodology: What are the desiderata for a biologically plausible implementation?)

- That it works, with reasonable parameters and within realistic ***brain*** time
- That it is in rough agreement with the neurolinguistic ***consensus*** on the function, if there is one
- Should be plausibly created by ***development***
- It is a proof of existence, not a hypothesis or argument that the function happens this ***exact*** way in the brain
- Multiple ***alternative*** implementations are a plus

Q: so, is this the neural basis of language?



A: not really, because
the hard part is filling the blanks!

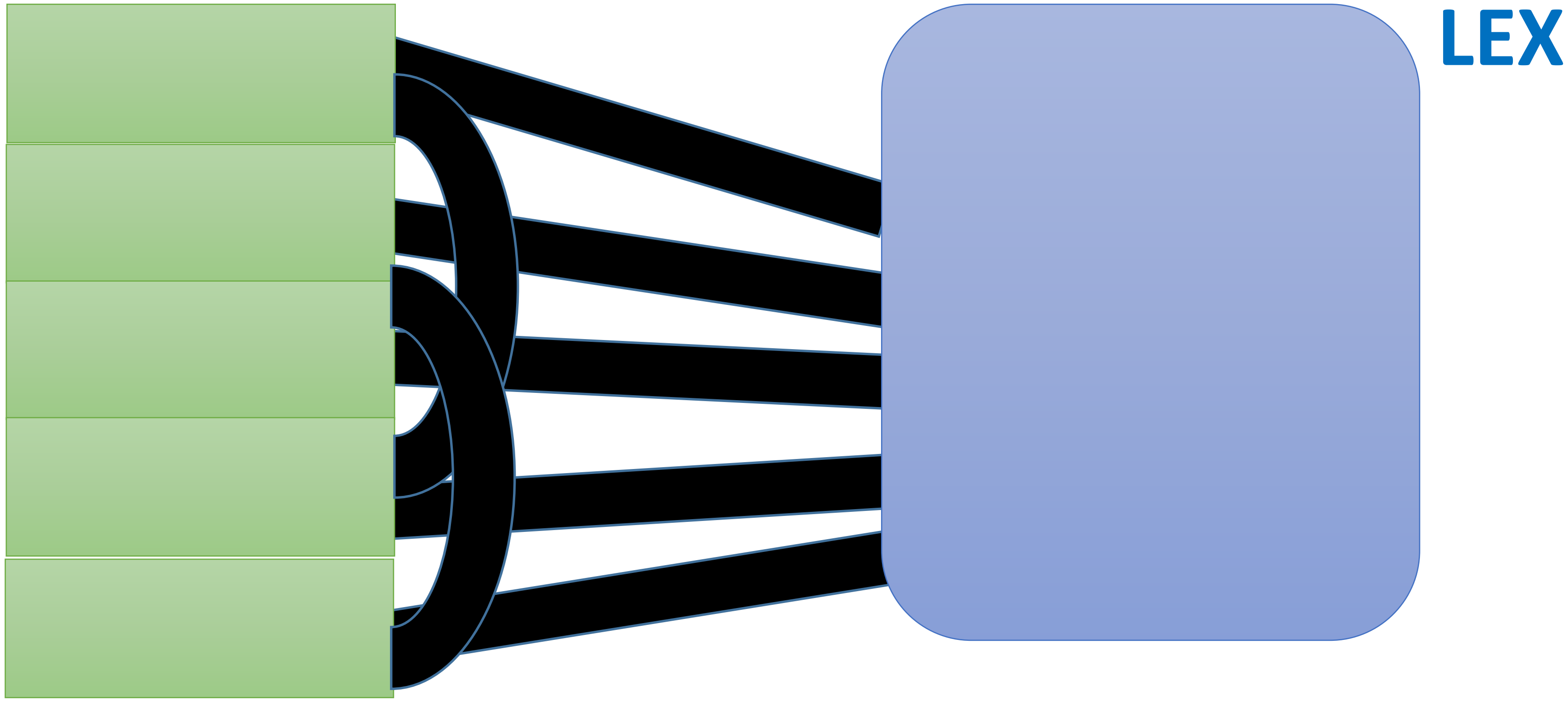


And how about *semantics*
and *generation*?

The neural basis of language?

This,
plus hardware for
semantics and
generation…

…plus the ability to learn all these things…



Current work: Biologically plausible language acquisition

Hardware: a tabula rasa of a couple of dozen brain areas, fibers, and populations of LRIs

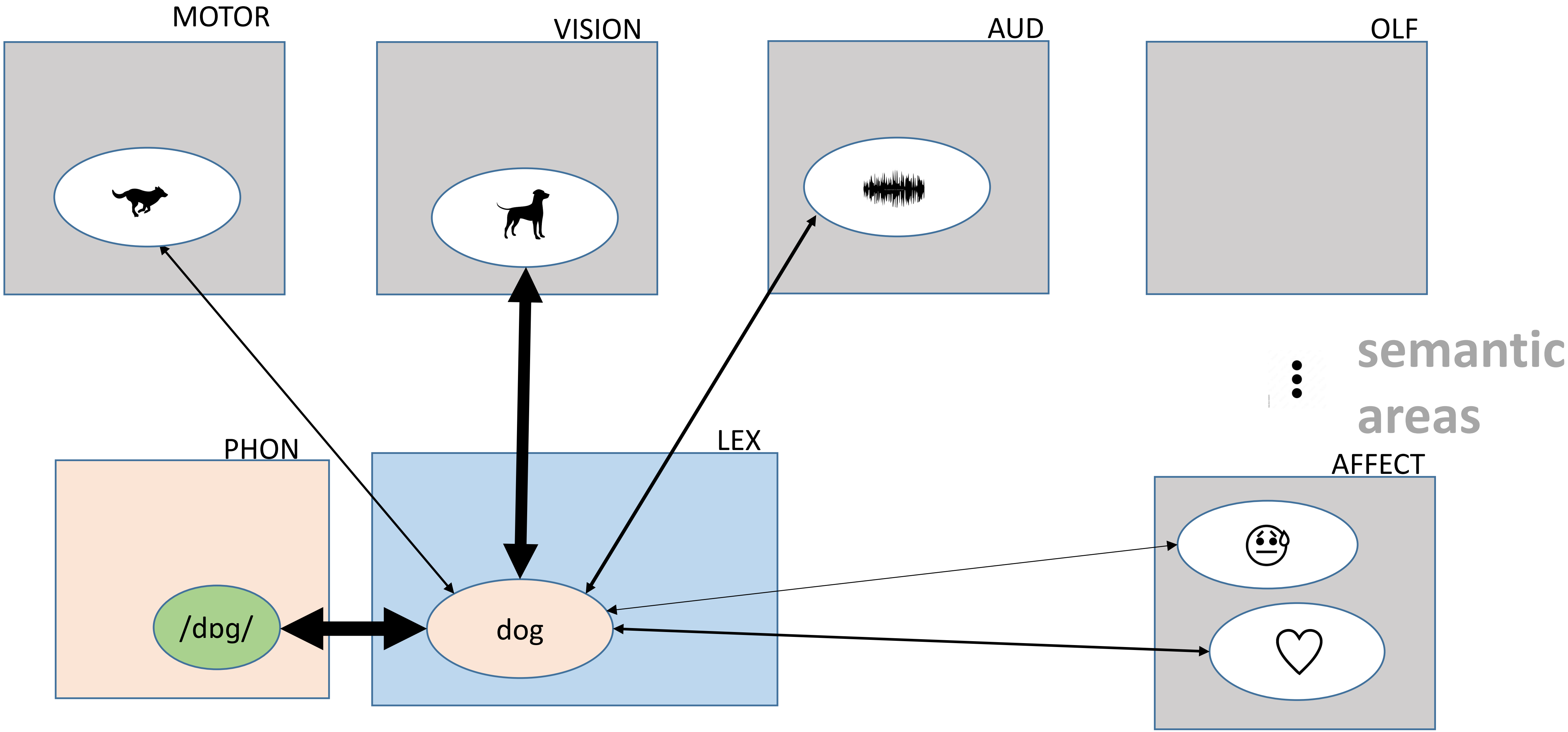
Input: a modest amount of ***grounded, shared-attention*** language

Output: a mature language organ, complete with a parser and a generator

Biologically plausible language acquisition: what we have implemented and tested so far

- Learning of nouns
- (We assume, as with the Parser, that ***phonetics*** has been solved)
- Representation of noun semantics
- Representation of verbs and their semantics
- Learning to generate two-word sentences
- Learning the word order of the language (two alternative ways)
- The **hard case:** learning syntactic roles of words from overheard language – still grounded and with shared attention.

Representation of nouns and their semantics



Representation of word semantics (cont.)

- Implements the “hub and spokes” semantics theory in neurolinguistics
- Each word representation consists of a ***star of assemblies***
- *Not unlike word2vec*
- The assemblies, their overlap, and the strengths of their connections are all **dynamic**
- They reflect cooccurrences of words, statistical regularities, order statistics, and changing world contexts

Representation of word semantics (cont.)

- In fact, *“man is to woman as king is to what?”* is answered ***automatically***
- “man” in LEX is connected with assemblies “human” and “male” in the semantic areas, “queen” is connected with “royal” and “female”
- *“I saw the tree with my binoculars”* will be disambiguated correctly
- “see” and “binoculars” will have strong overlap in the semantic areas, and as a result the prepositional phrase will synapse preferentially with the verb

…and of course,
no matter how clever you are,
you could be doing
the wrong thing…

Sooooooo…

- The study of the brain is fascinating and bottomless
- Language is the hardest thing any brain has done, and so it must hold the key
- Implementing language in the brain entails close encounters with profound questions
- Progress on biologically plausible acquisition
- **A promising way forward in AI?**
- **Are assemblies and long-range interneurons the seat of Axel’s “logic”?**

my collaborators!

Santosh
Vempala
GaTech

Max Dabagia
GaTech

Mike Collins
Google

Dan
Mitropolsky
Columbia

Pilu Crescenzi
Gran Sasso

Merci!