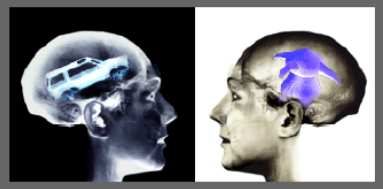


UCSD

Temporal Dynamics of Learning Center

How does the Brain make the Mind?

Gary Cottrell

Computer Science and Engineering Department

Institute for Neural Computation

Temporal Dynamics of Learning Center

UCSD

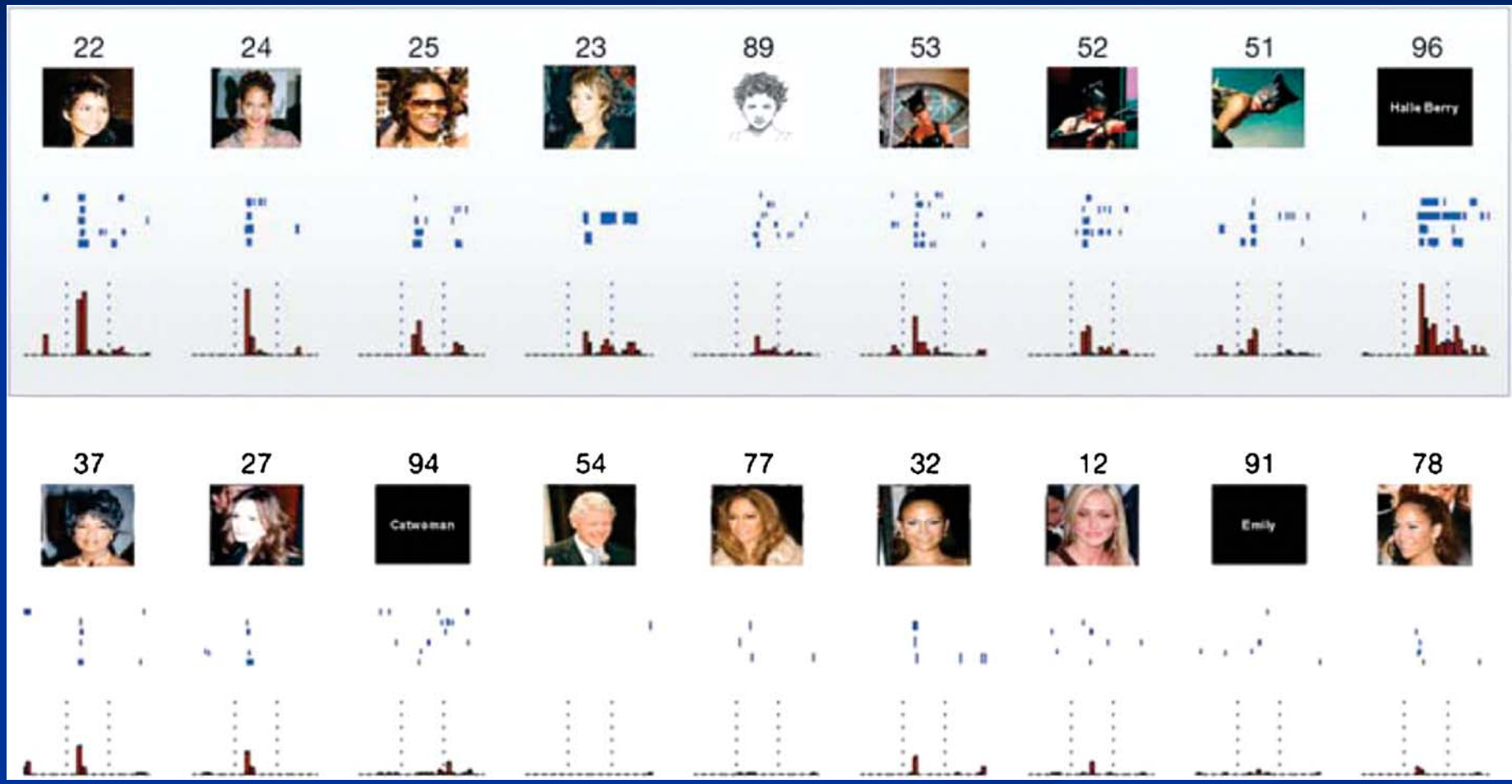
Introduction

- Your brain is made up of 10^{11}-10^{12} neurons - brain cells - that communicate over 10^{14}-10^{15} connections.
- Each cell only "sees" its input from other cells - and only sends electrical spikes to other cells.
- How in the world could this device make a mind???

Introduction

- There are many ways to try to figure out how this beast works:
  - Measure *behavior* - getting an idea of the way the brain takes input and turns it into output.
  - Measure *brain waves* using electrode caps (EEG - ElectroEncephaloGram)
  - Measure *brain activation* using giant magnets (fMRI - functional Magnetic Resonance Imaging)
  - *Record neurons* in the brains of other animals while they are performing some task
  - *Record neurons* in the brains of human patients getting surgery for epilepsy

The "Halle Berry" neuron…



Recorded from a human subject awaiting brain surgery…

The main idea of this talk

- One way to try to understand how this thing works is to build a *working model* of it - using data from the other approaches.
- Nowadays these models can be run on computers - so that we can actually "see" them work.
- This is the role of *cognitive science* - or to put it more specifically: *computational cognitive neuroscience*

The three axioms of cognitive science (a la Cottrell)

1. The mind is what the brain does
   - There is no "spooky stuff"
2. What the brain does, i.e., *thinking*, is a kind of computation
   - Love at first sight is a computation
3. The kind of computation the brain does is *probabilistic* - probability is the "language of thought"
   - It must be, in order to deal with the uncertain nature of the world

Neural networks are a good model of this kind of computation

What is a "working model"?

The "visible V-8" by Revelle

What is a "working model"?

The "visible V-8" by Revelle

What is a "working model"?

- Similarly, our models won't replace your brain - yet!
- But because they are made on a computer, they can *do* things - like recognize faces, interpret sentences, move limbs, play games, drive cars

Outline

I. Motivation: Why neural nets?

II. Human-style computation: What's it like?

III. Neural nets: What are they like?

The Interactive Activation Model of Reading

IV. Conclusions

Motivation

- Why are people (still) smarter than machines?
- Is it because we have faster hardware?
  - No
- Is it because we have better programs?
  - No
- It's because we have brains! ;-)
  - Brains are massively parallel machines

Motivation

- Why are people (still) smarter than machines?
  - Basic differences in architecture
- Hence we study *brain-like* computational models: Neural nets
- "Cartoon versions" of how the brain works

Motivation

- And now we can *train* neural nets to do things we want them to do.
- So *learning* is a big plus
- Recent development: "deep networks" - we can now train networks with many layers of processing - like our brains have!

Outline

I. Motivation: Why neural nets?

II. Human-style computation: What's it like?

III. Neural nets: What are they like?

The Interactive Activation Model of Reading

IV. Conclusions

Combining Information

- People are able to combine *lots* of different kinds of knowledge *quickly* in understanding English -
- For example, in understanding the relationships given in a sentence:
  - *Syntax* (structure) gives us some information:

    The boy kissed the girl.

Combining Information

- But usually we need *semantics* (meaning) too:

  I saw the man on the hill with the big hat.

  I saw the man on the hill with my telescope.

- These have identical parts:

  [Pronoun] [Verb] [Noun Phrase] [Prep. Phrase] [Prep. Phrase]

- But how those parts go together depends on the meanings…

  I saw [the man [on the hill] [with the big hat]].

  I saw [the man [on the hill]] with my telescope.

Combining Information

Ditto for pronoun reference:

"The city council refused the demonstrators a permit because ***they*** were communists."

Who are ***they***?

Combining Information

"The city council refused the demonstrators a permit because ***they*** were communists."

Who are ***they***?

In San Diego, it is the ***demonstrators*** who are likely to be the communists

Combining Information

"The city council refused the demonstrators a permit because ***they*** were communists."

Who are ***they***?

In Berkeley, it is the ***city council*** who are likely to be the communists! ;-)

Combining Information:
Word Sense Disambiguation

The most frequent words (in any language) are the most ***ambiguous!***

How do we deal with this ambiguity?

Answer: We combine constraints from many sources…

Combining Information:
Word Sense Disambiguation

Things that people do well involve integrating constraints from multiple sources:

Discourse Context: "I'm ***late***"

Grammar: "The carpenter ***saw*** the wood"

Meaning frequency: "Bob ***threw a ball***"

Combining Information:
Word Sense Disambiguation

Semantics (meaning):

Different agents can change meaning:

"Billy picked up ***the truck*** and threw it across the room."

"Superman picked up ***the truck*** and threw it across the street."

Associations between word senses:

"dog's ***bark***" "***deep pit***"

Combining Information:
Word Sense Disambiguation

Common sense knowledge:

"The man walked on the ***deck***"

"Nadia swung the hammer at the nail and ***the head*** flew off" (Hirst, 1983)

Computers are different…

They are fine with sentences like:

The boy the girl the dog bit liked cried.

You don't think that's a sentence?

Computers are different…

The boy cried.

The boy the girl liked cried.

The boy the girl the dog bit liked cried.

Computers are different…

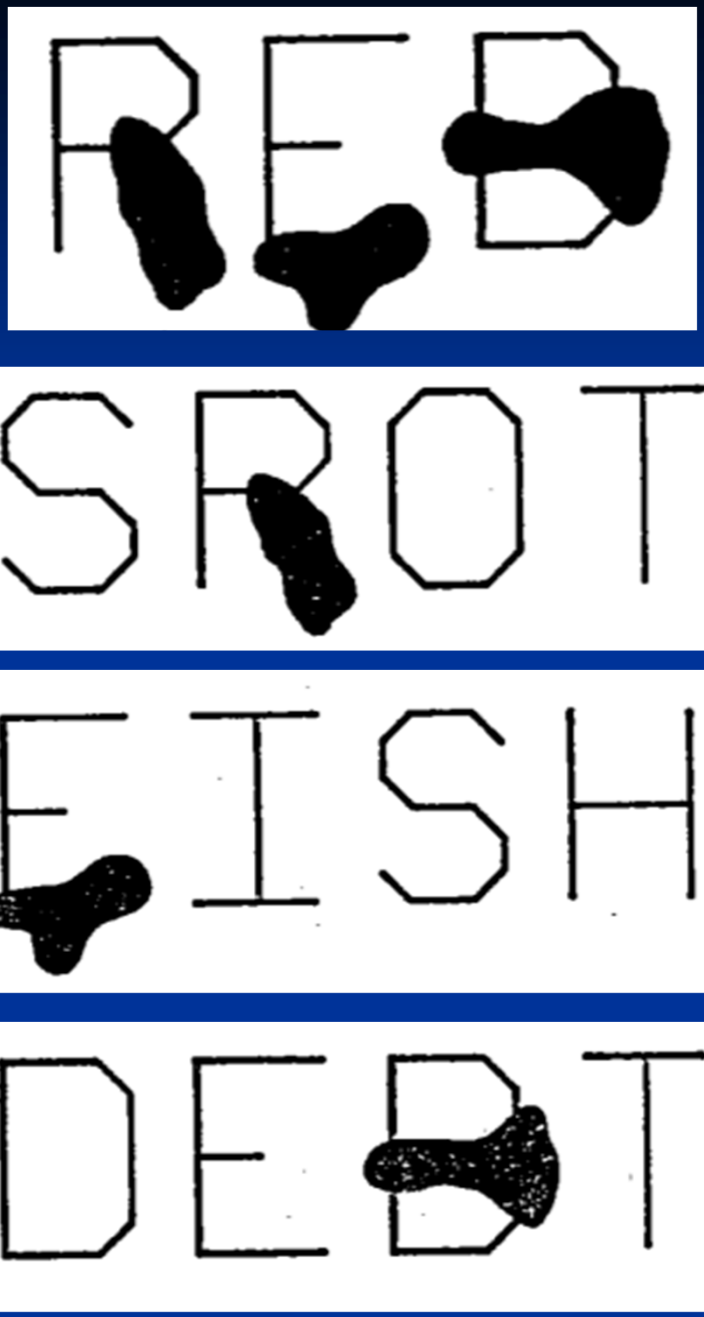
But…

U CN REED THIIS CANDT YU?

Mutual Constraints: Audience Participation! Read this aloud with me:

Read these aloud with me:

But…



What do you see?
Giuseppe Arcimboldo(1527-1593)

What do you see?
Giuseppe Arcimboldo(1527-1593)

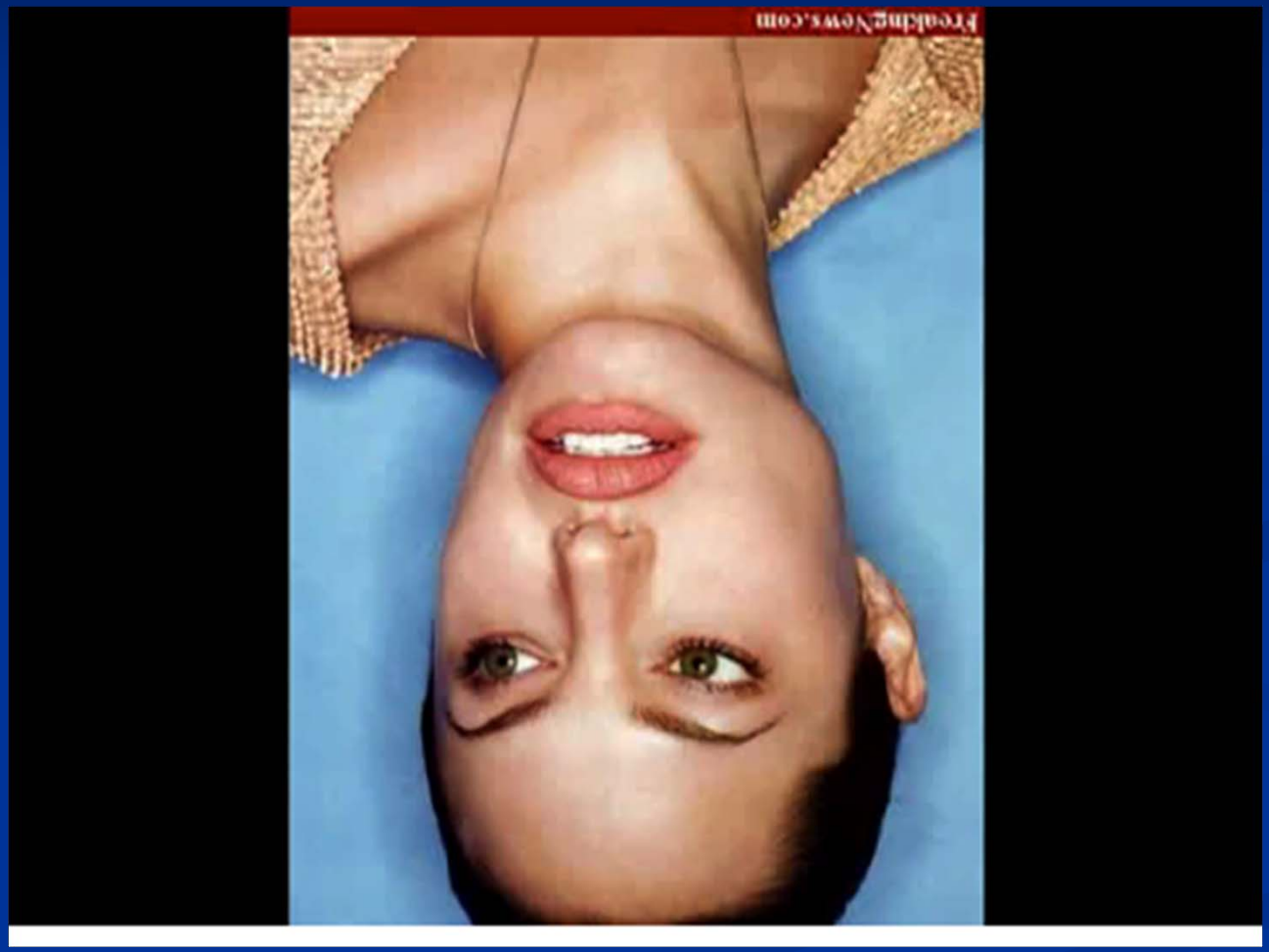
Who do you see?

Who do you see?

Who do you see?

- Context influences perception

Who do you see?



How do humans compute?

- They are fast at combining information to:
  - Understand sentences
  - Disambiguate words
  - Read ambiguous letters
  - Recognize faces (and sometimes the information leads us astray!)

Outline

I. Motivation: Why neural nets?

II. Human-style computation: What's it like?

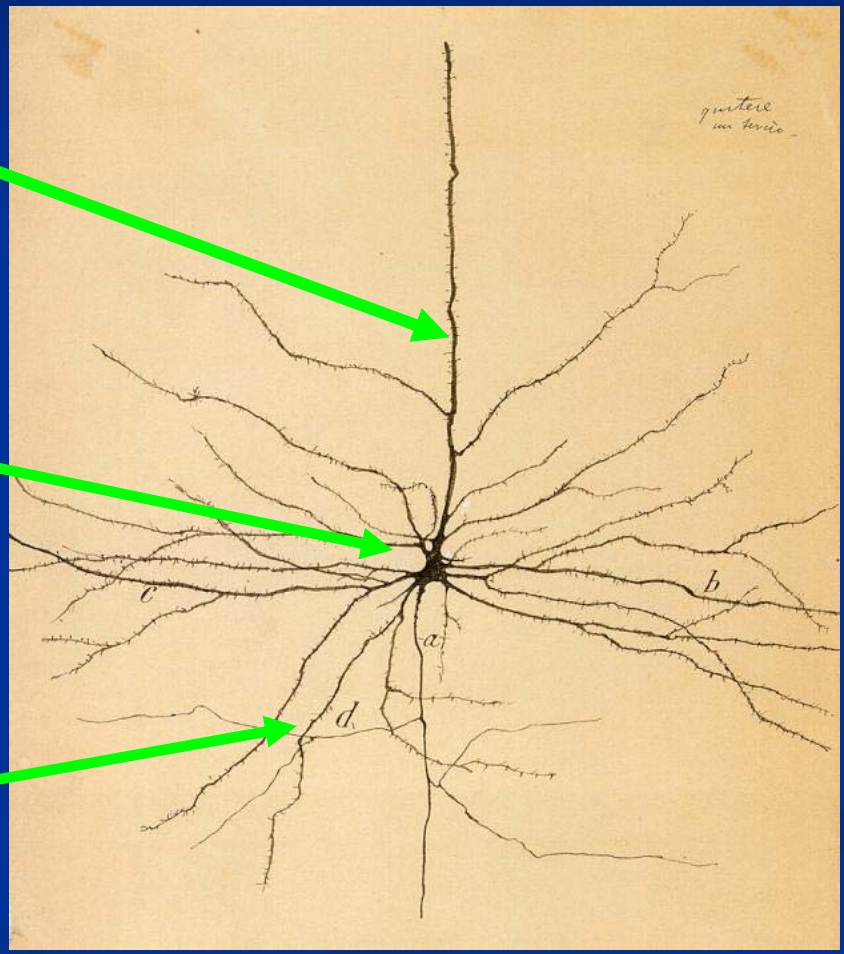
III. Neural nets: What are they like?

The Interactive Activation Model of Reading

IV. Conclusions

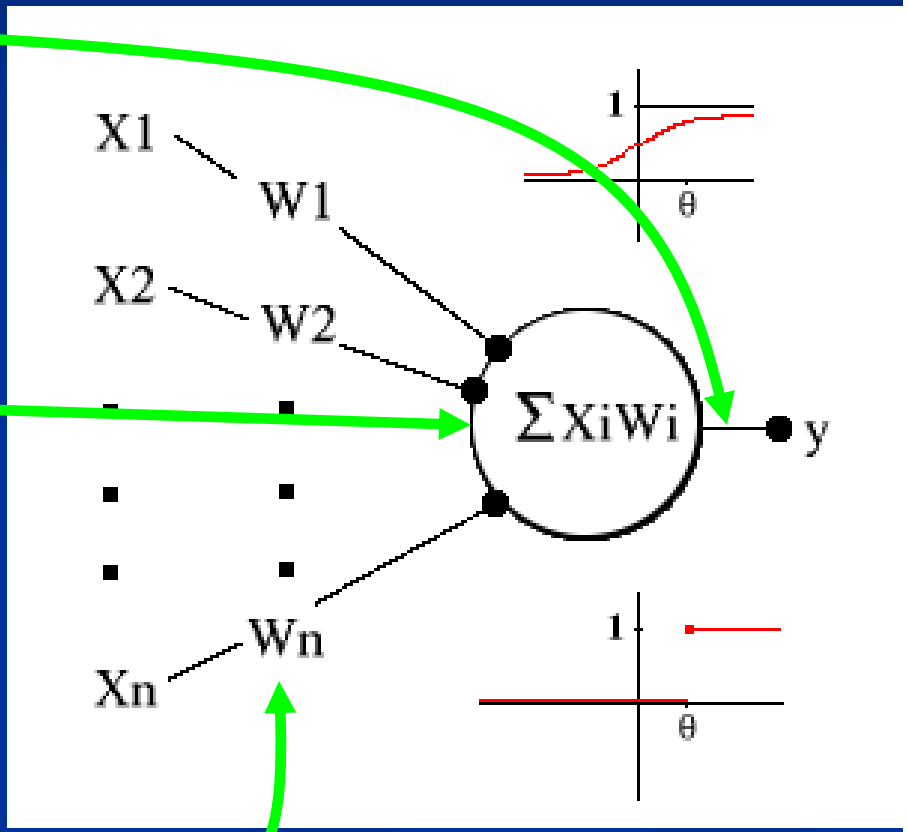
A real neuron
(Ramon Y Cajal, 1899)

- Axon (output)
- Cell body (soma)
- Dendrites (input)



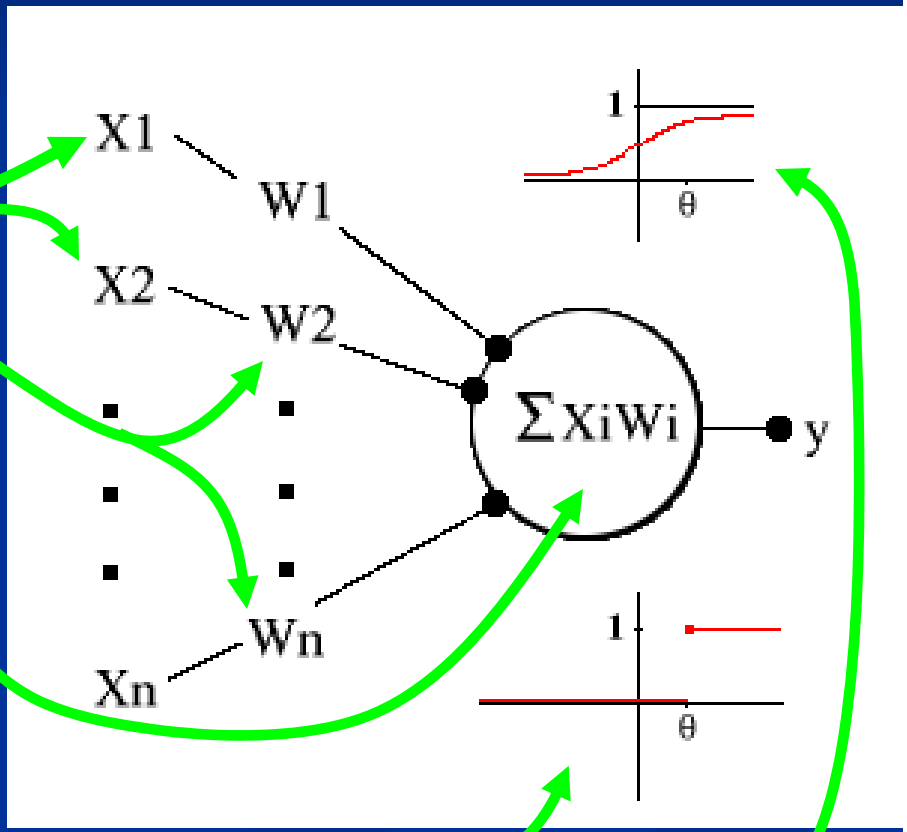
A model "neuron"

- Axon (output)
- Cell body (soma)
- Dendrites (input)



A model "neuron"

- Inputs (from another unit or the outside world)
- Connection strengths (or *weights*)
- Internal "potential"
- Output, representing firing frequency

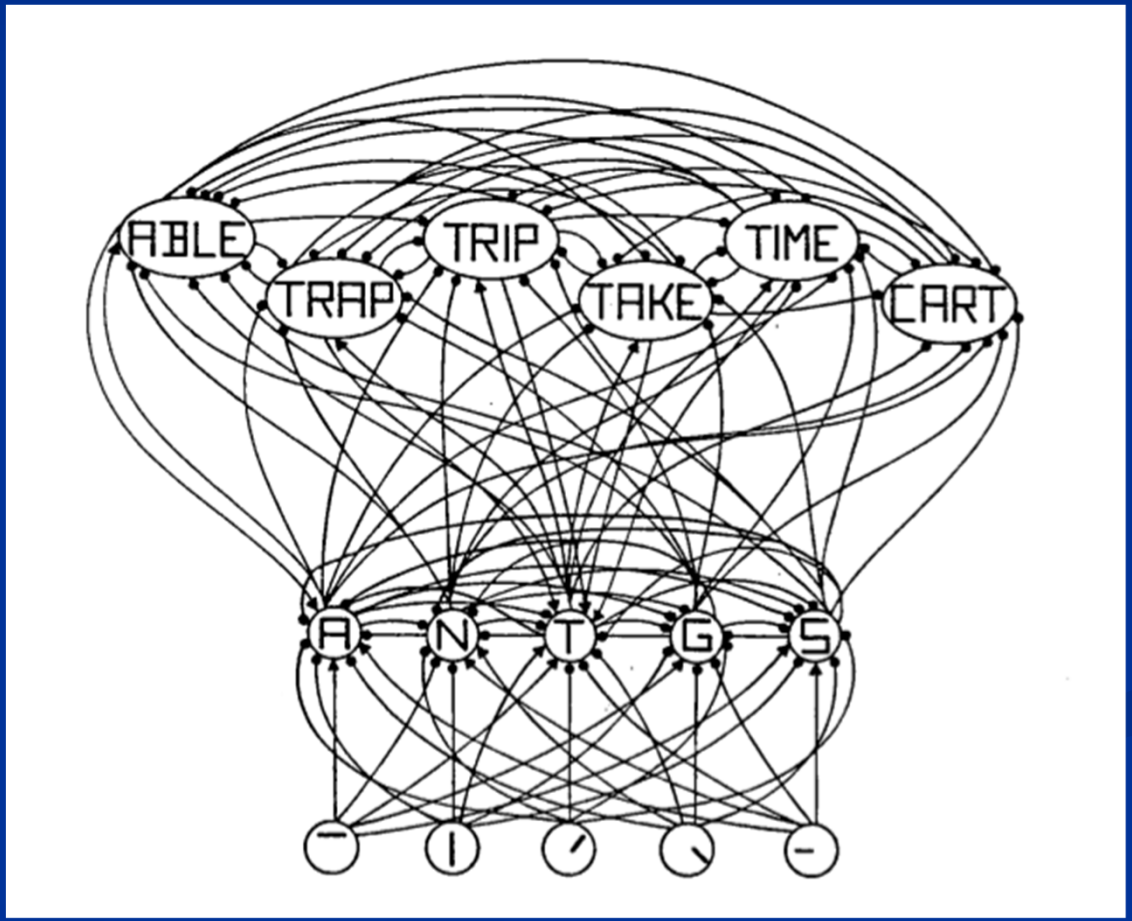


Neural nets
(PDP nets, connectionist networks)

- ***Networks*** of simple units
- Connected by ***weighted links***
- Compute by ***spreading activation and inhibition***

Your first neural net: The Interactive Activation Model:
A model of reading from print

- Word level
- Letter level
- Feature level



Your first neural net: The Interactive Activation Model:
A model of reading from print

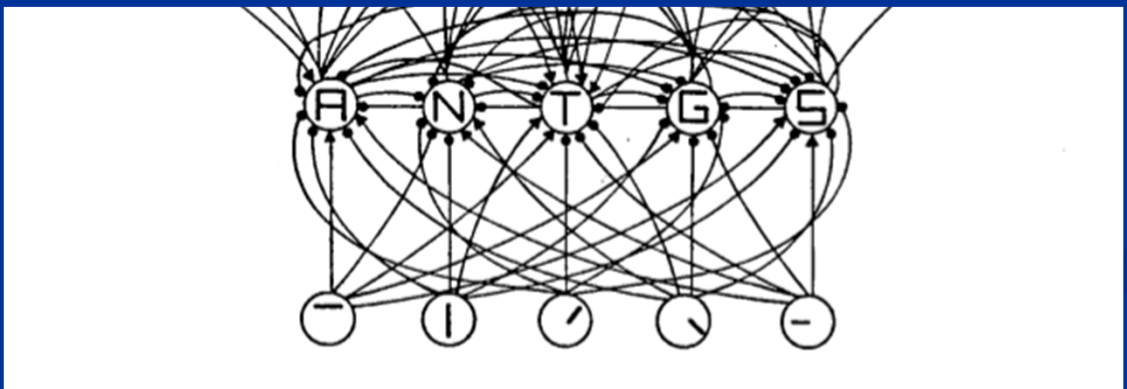
- Feature level

Your first neural net: The Interactive Activation Model: A model of reading from print

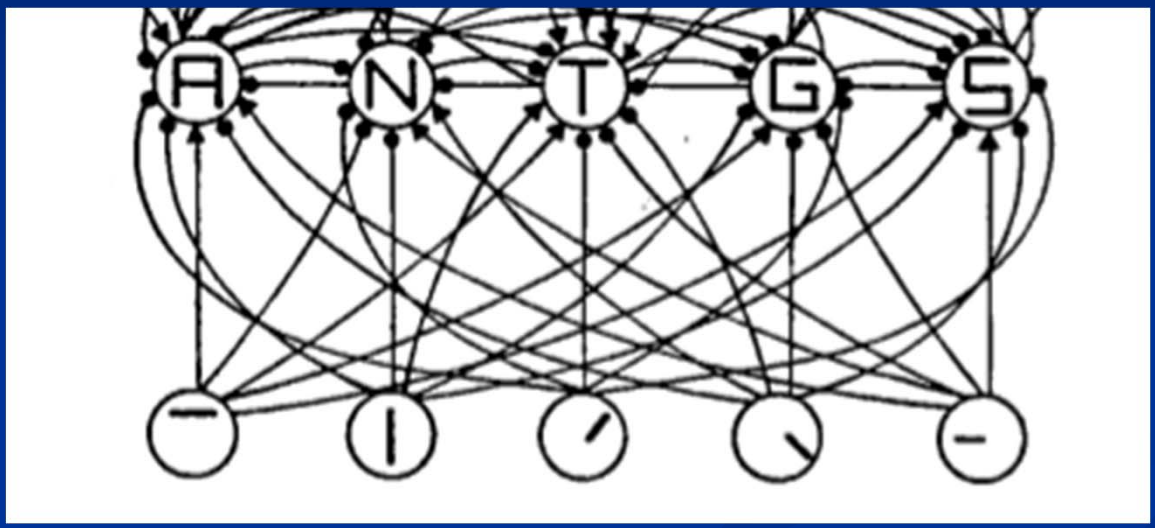
- Feature level:
  - Little *bar detectors* - that we know exist in the first stop in the visual part of your brain.
  - There are enough of these to represent every letter
  - These are copied to make bar detectors for each letter in a word.

Your first neural net: The Interactive Activation Model:
A model of reading from print

- Letter level
- Feature level



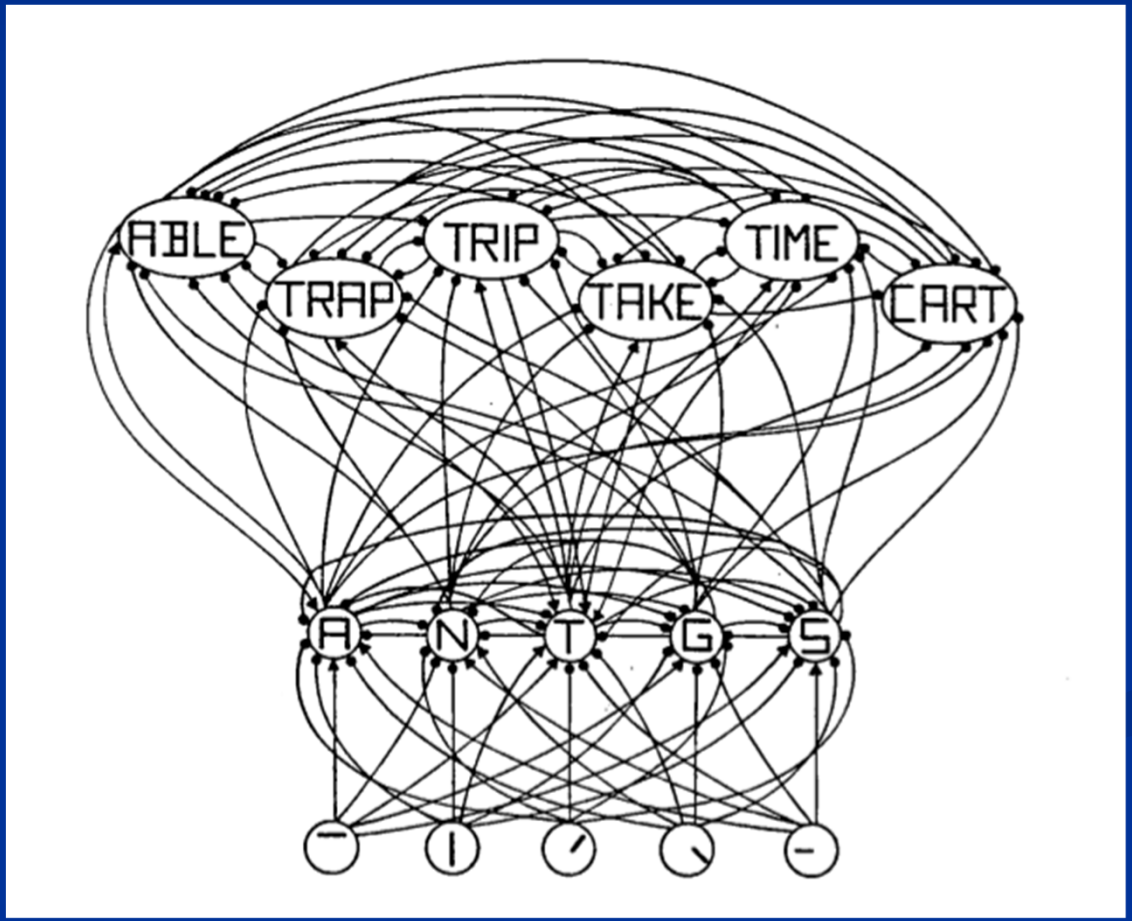
The Interactive Activation Model:
A model of reading from print



- The features excite compatible letters - and *inhibit* incompatible ones
- The little circles are inhibitory (negative) links
- The arrows are excitatory (positive) links
- The letters (in each position) fight it out through inhibitory links - because there can only be one letter in each position. This is called a "Winner Take All" network

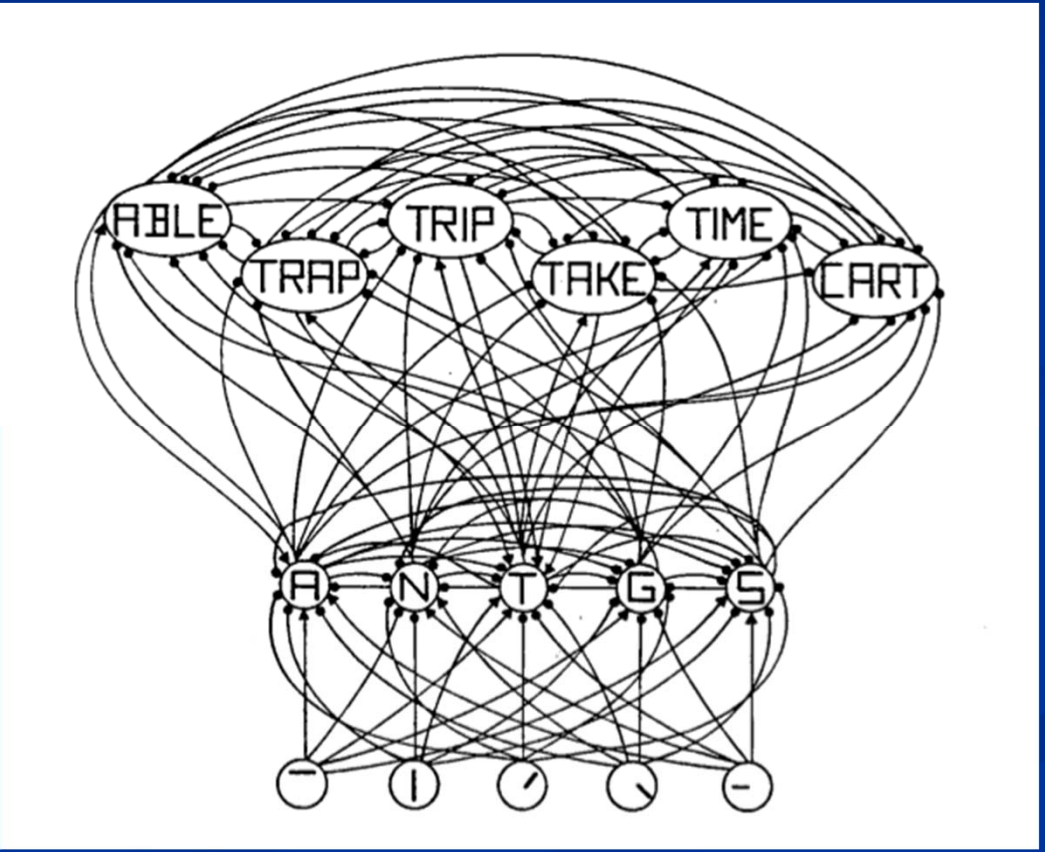
Your first neural net: The Interactive Activation Model:
A model of reading from print

- Word level
- Letter level
- Feature level



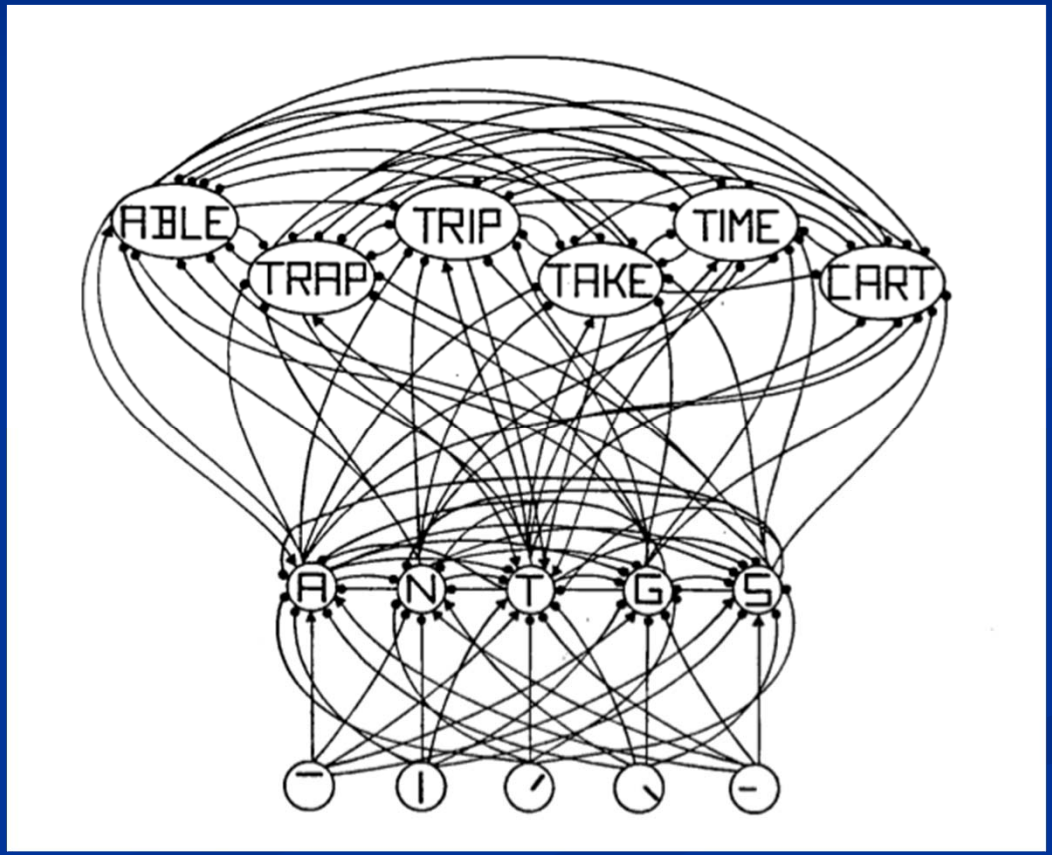
Your first neural net: The Interactive Activation Model:
A model of reading from print

- The word level units are activated by letters that are compatible.
- This picture shows the neighbors of the letter "T" in the first position.
- "T" votes *for TRIP, TRAP, TAKE* and *TIME*
- "T" votes *against CART* and *ABLE*
- *Quick Quiz*: Why does it vote against *CART*?

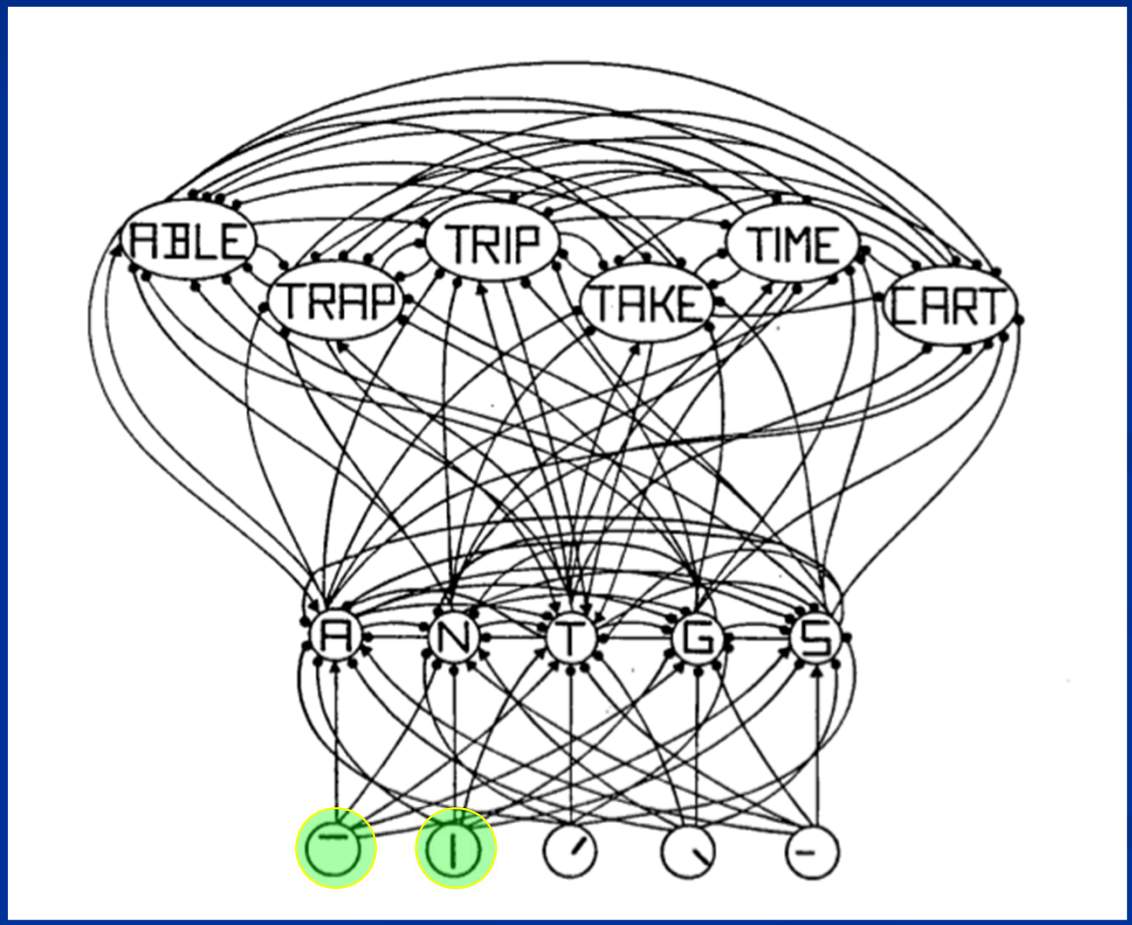


Your first neural net: The Interactive Activation Model:
A model of reading from print

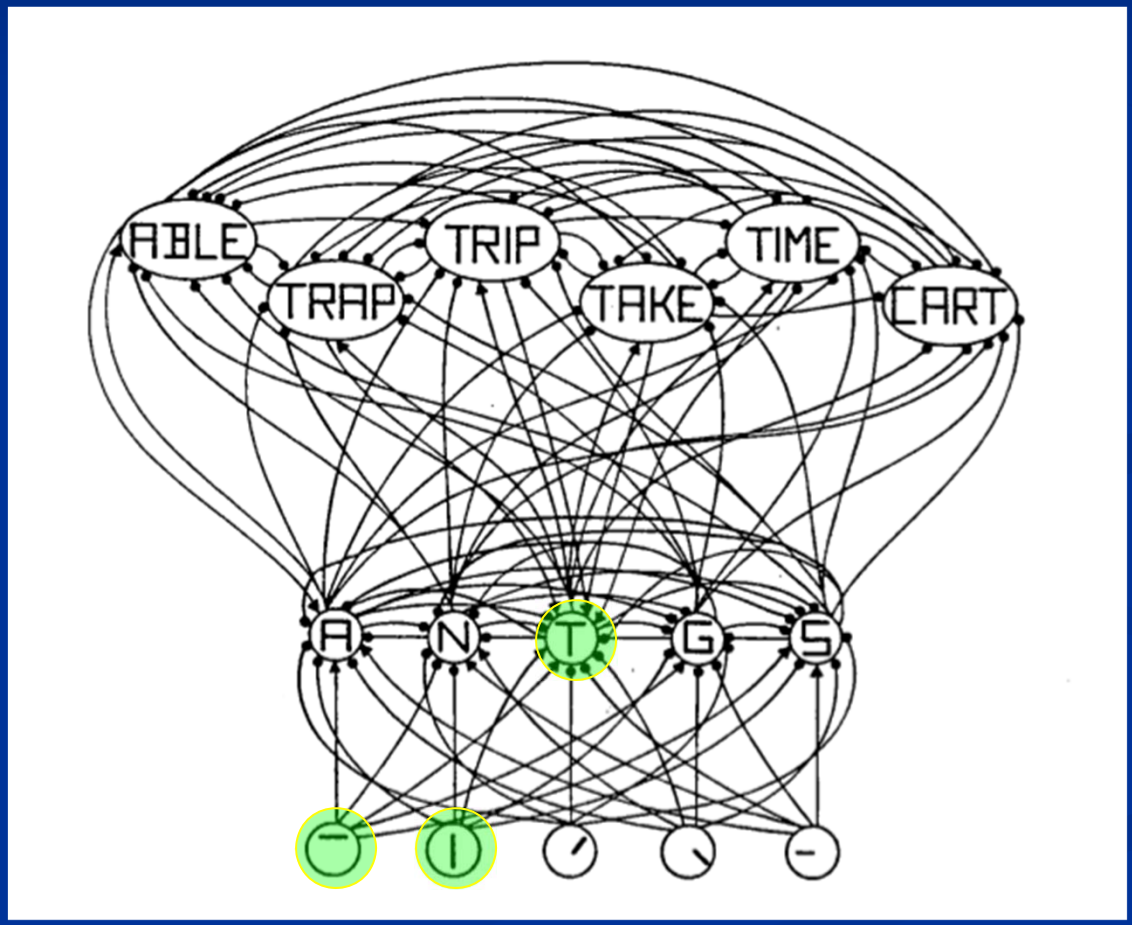
- Crucially in this model, word units feed back on the letter units, making them more active than they would be otherwise.
- This accounts for the "word superiority effect" - you are better at seeing letters when they are part of a word than when they are in a nonword letter string



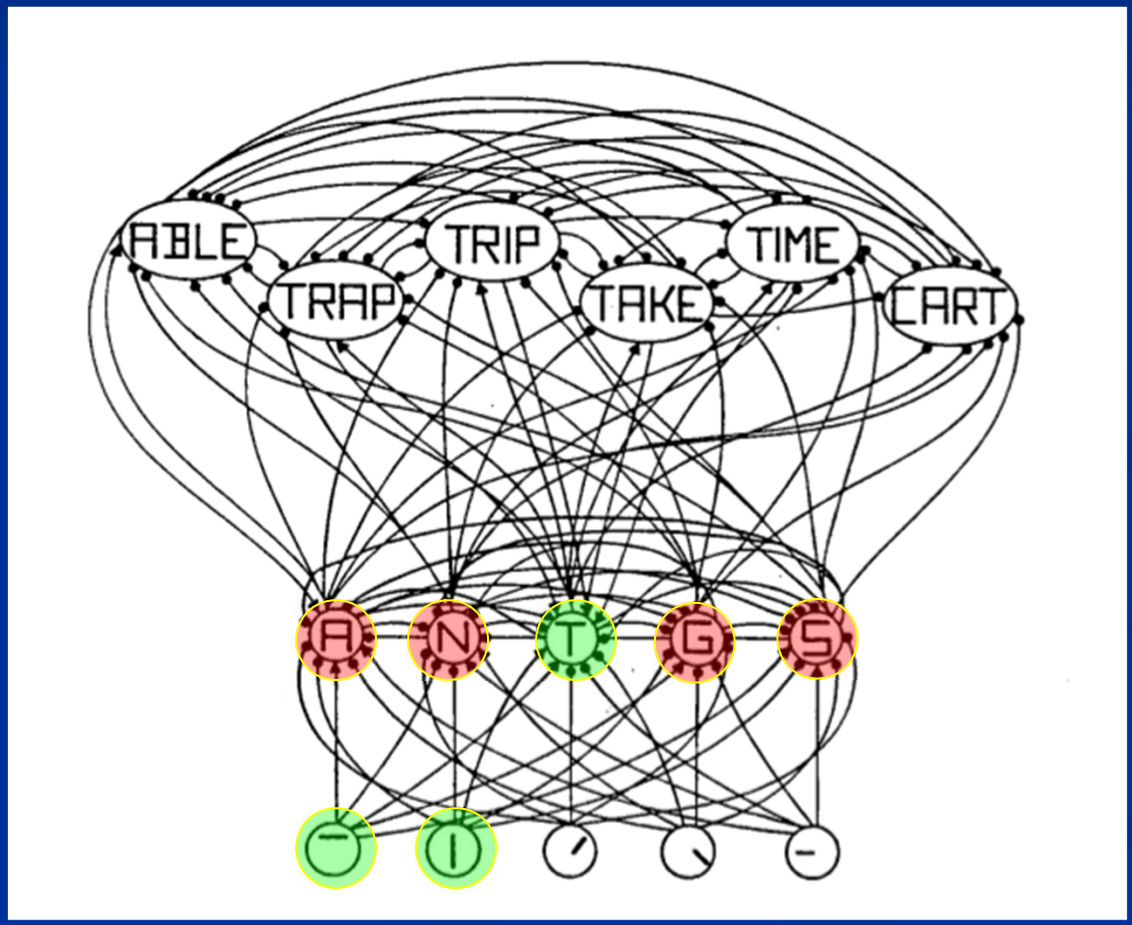
Activating "T"



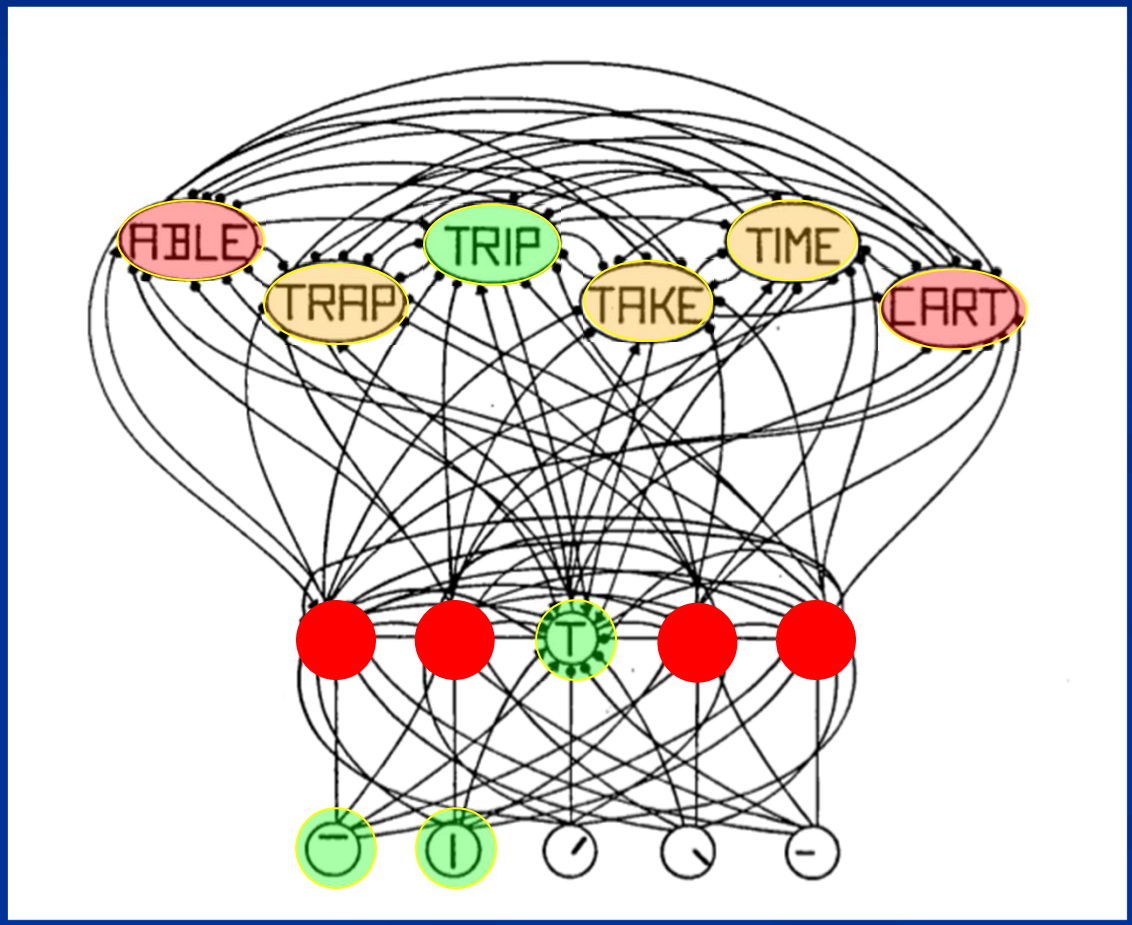
Activating "T"



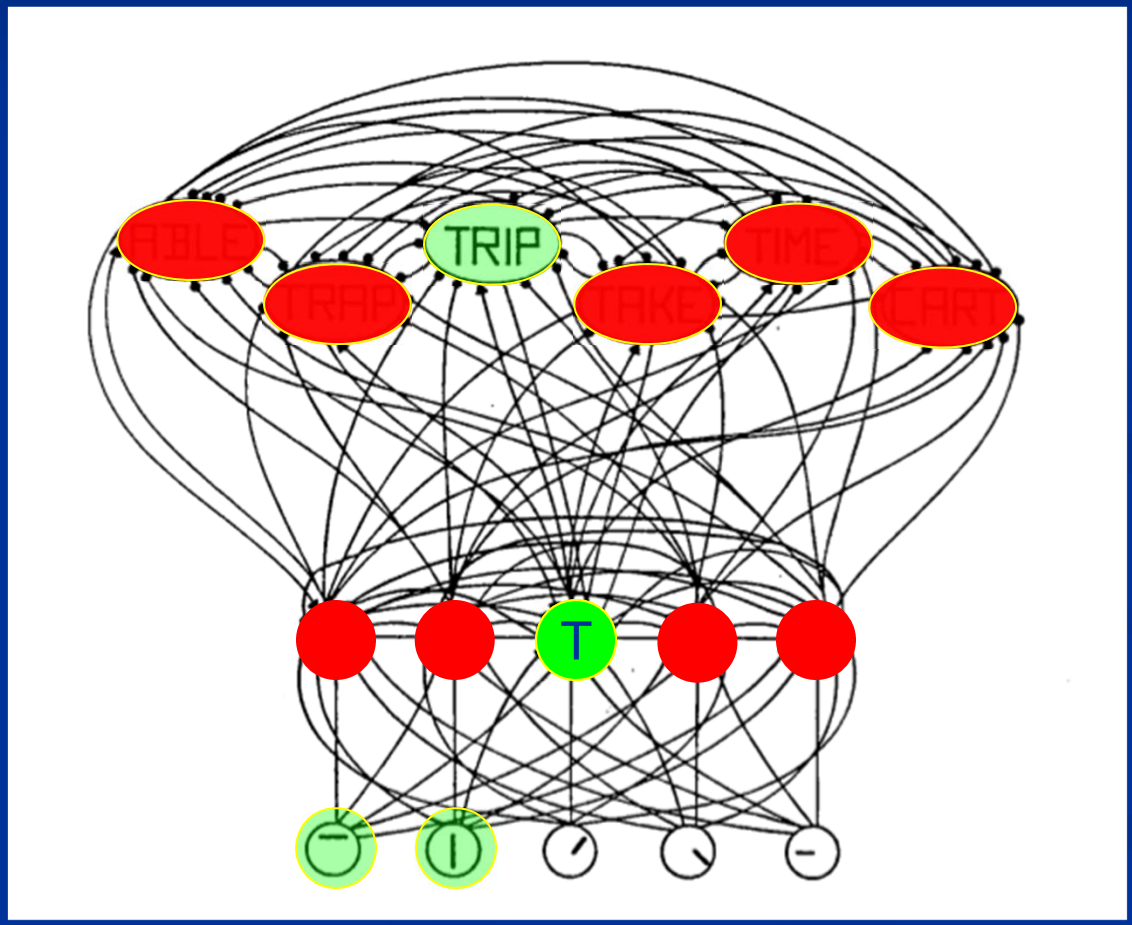
Activating "T"



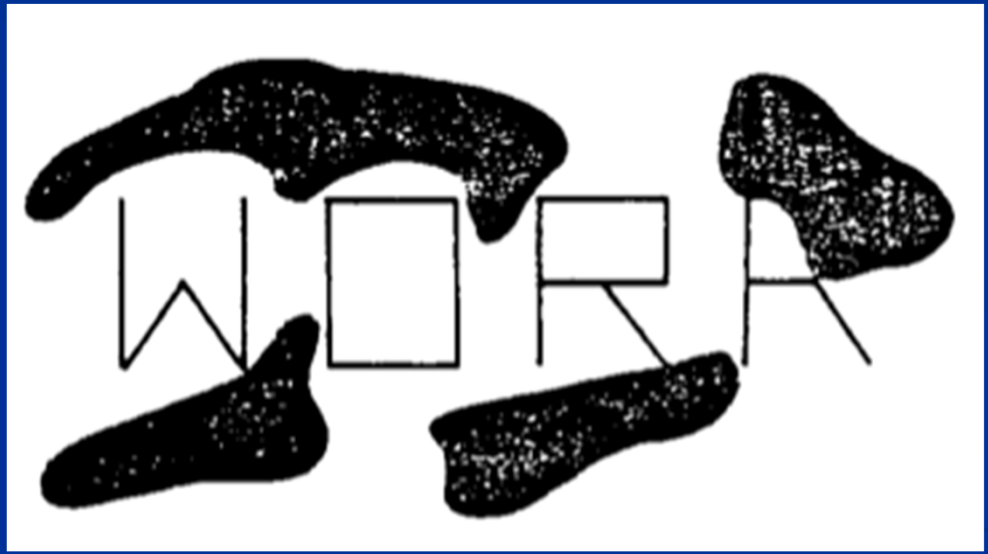
Activating "T"



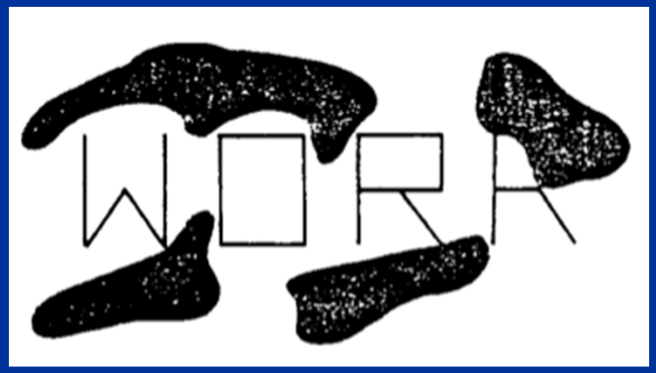
Activating "T"

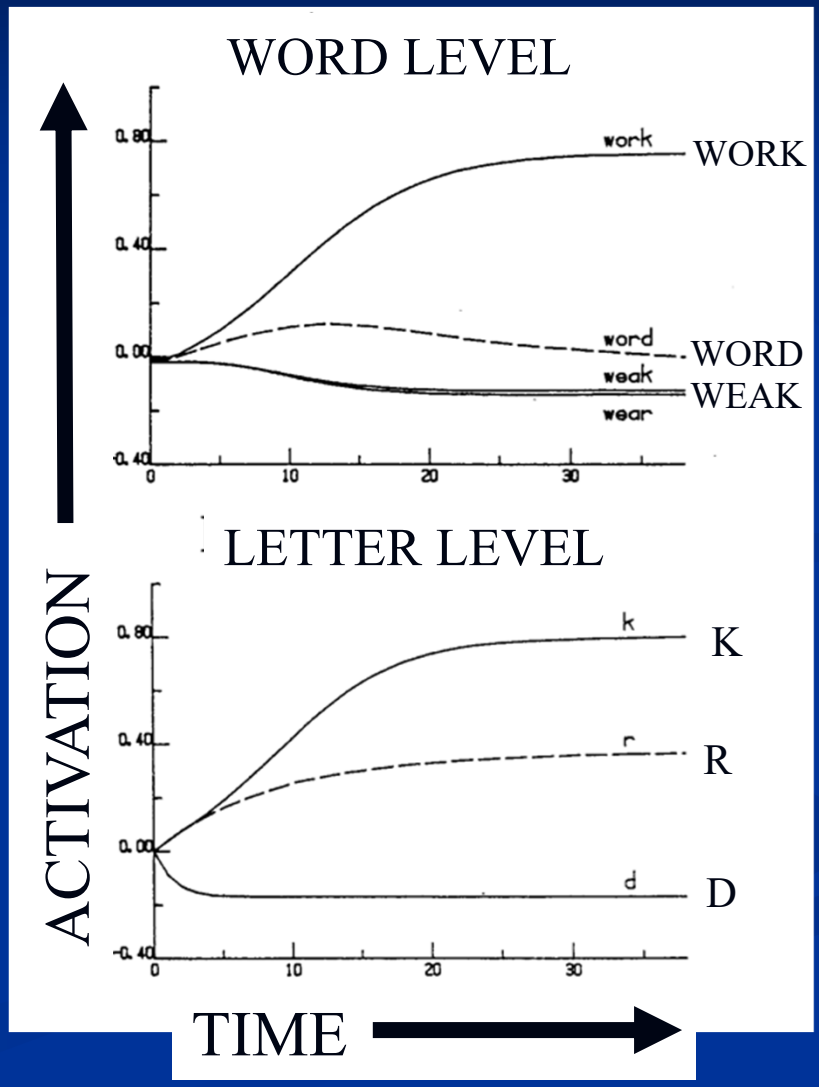


Read this word with me…

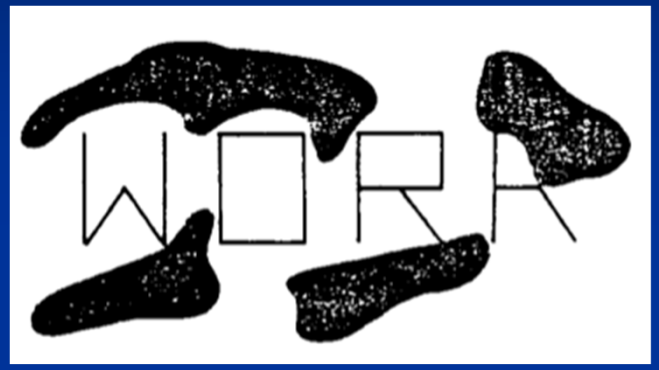


Operation of the model

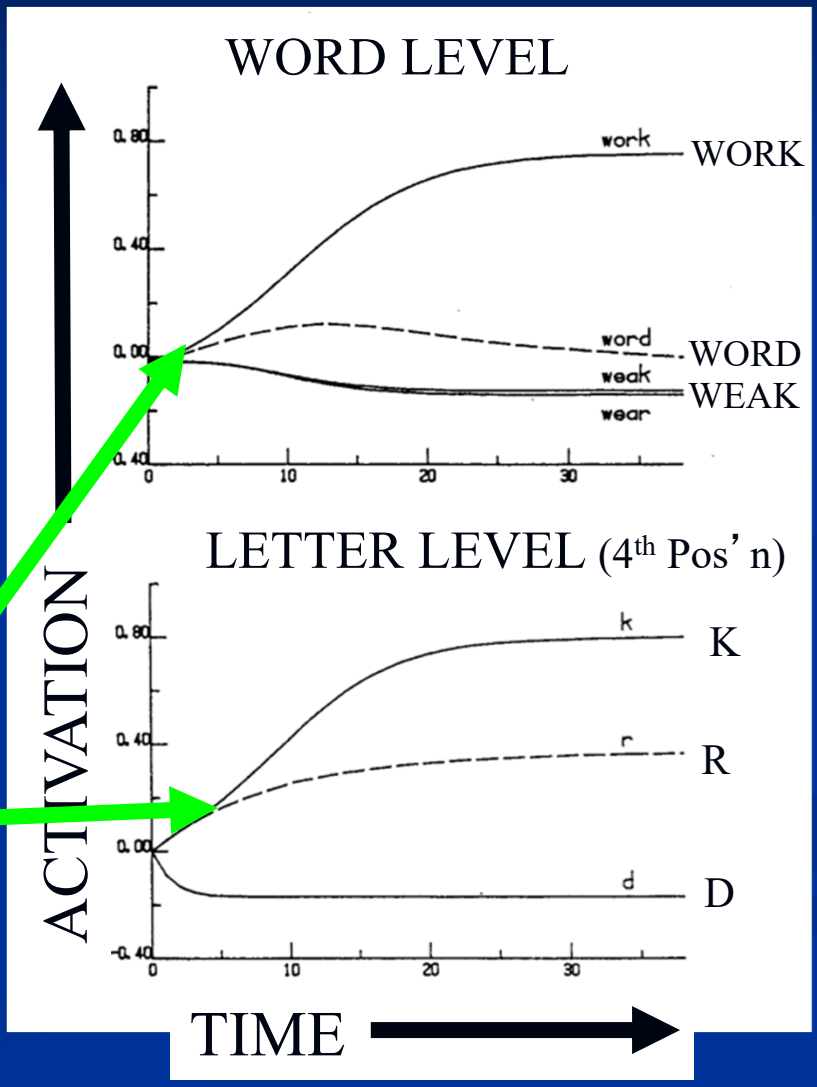




Operation of the model

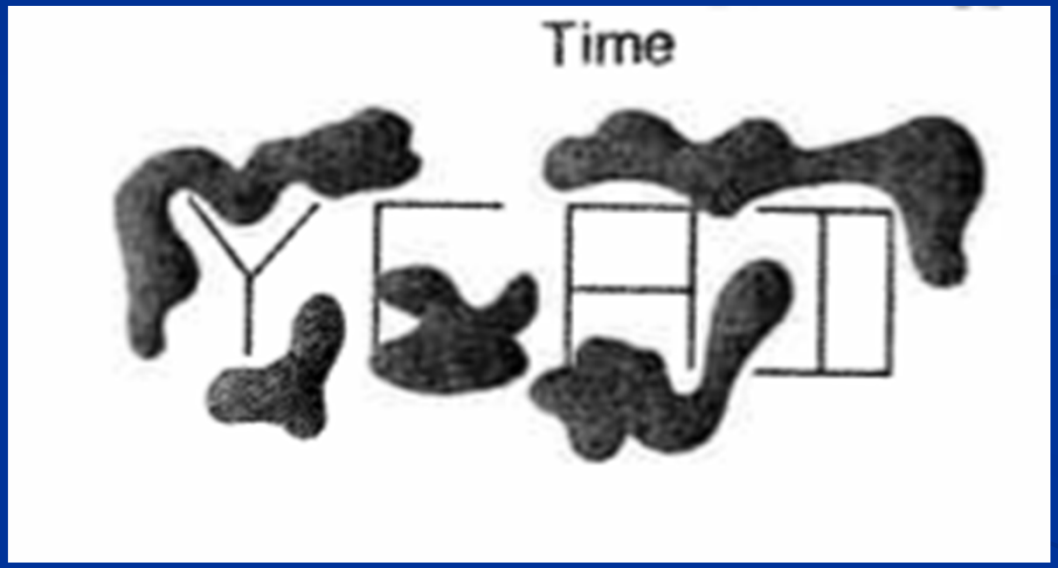


Note here:
WORK is starting to win out over WORD and WEAK, giving more feedback to K in position 4

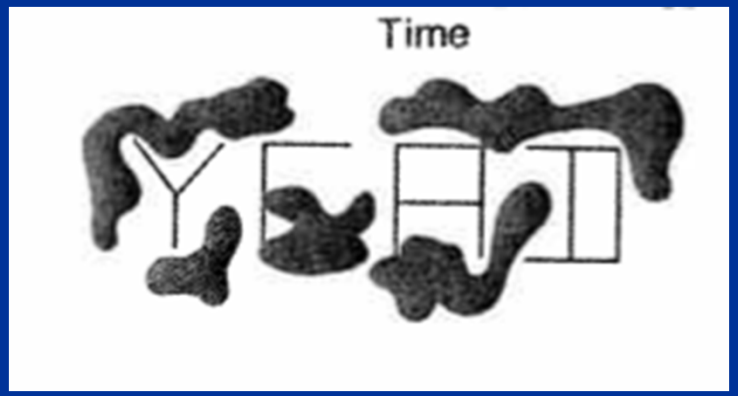


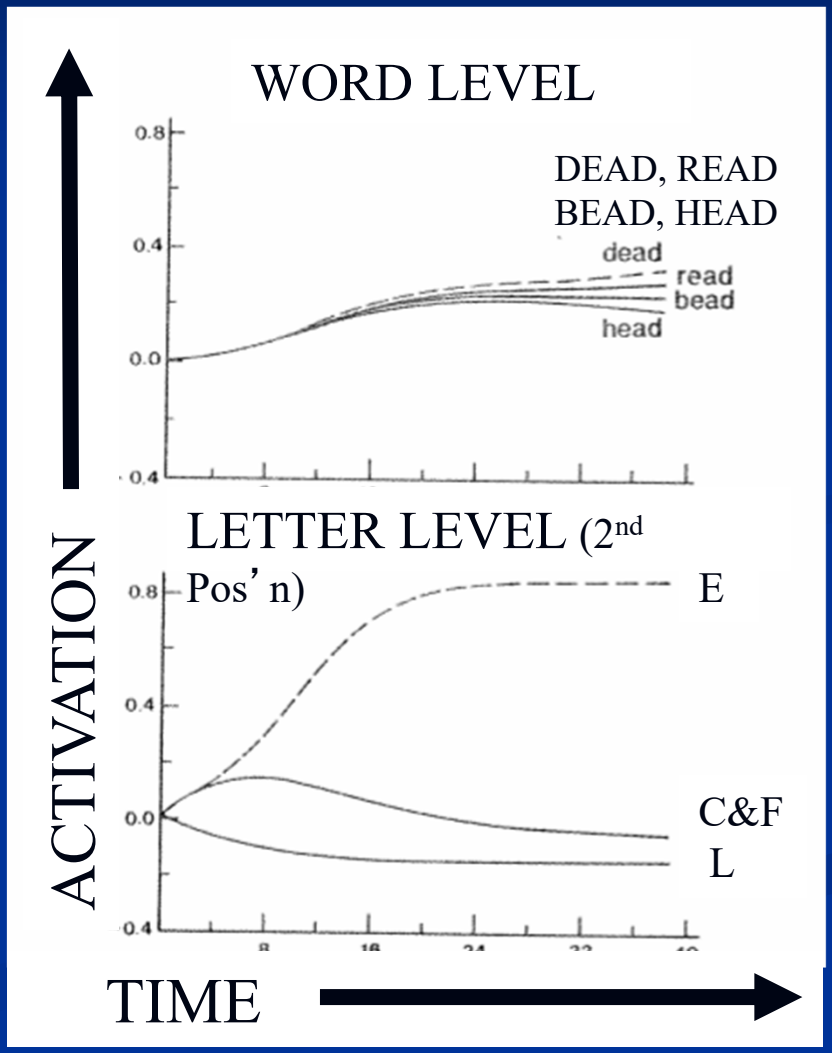
Example of data accounted for… Pseudoword effect

- Read this (pseudo-)word with me:



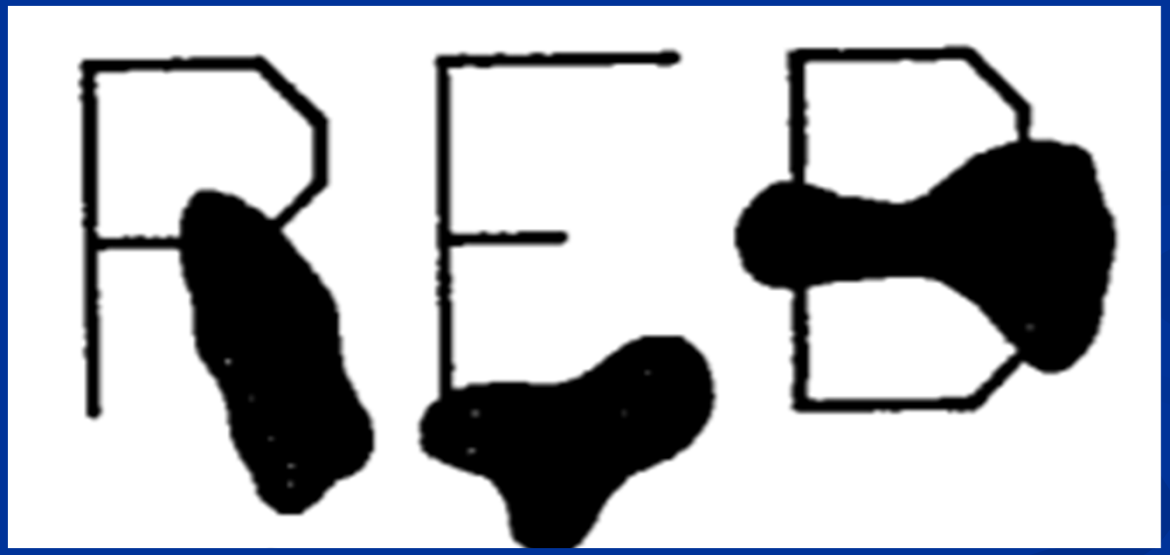
Example of data accounted for… Pseudoword effect



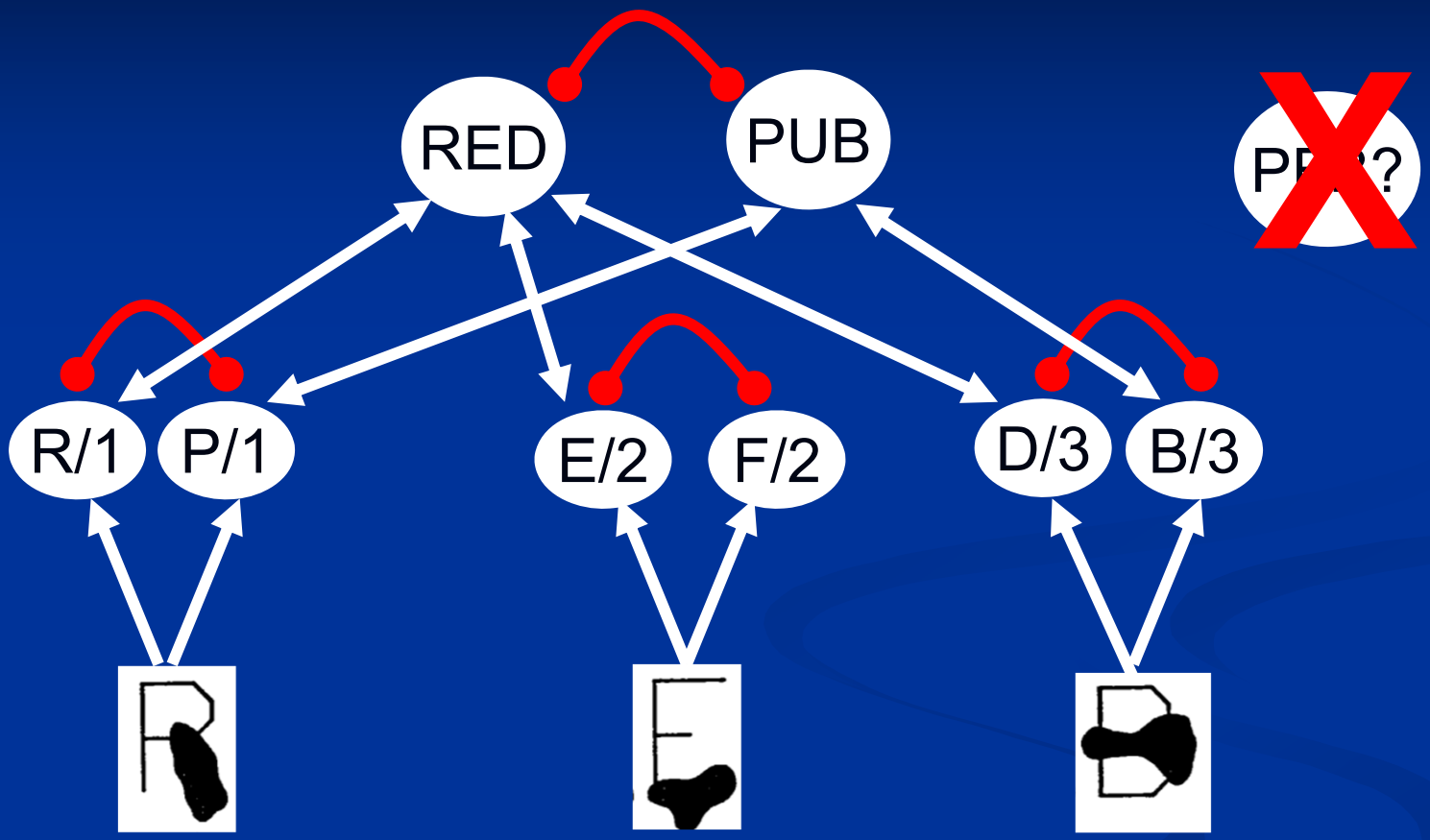


Operation of the model

- How does the model explain why we all read this as “RED” and not “PFB”???

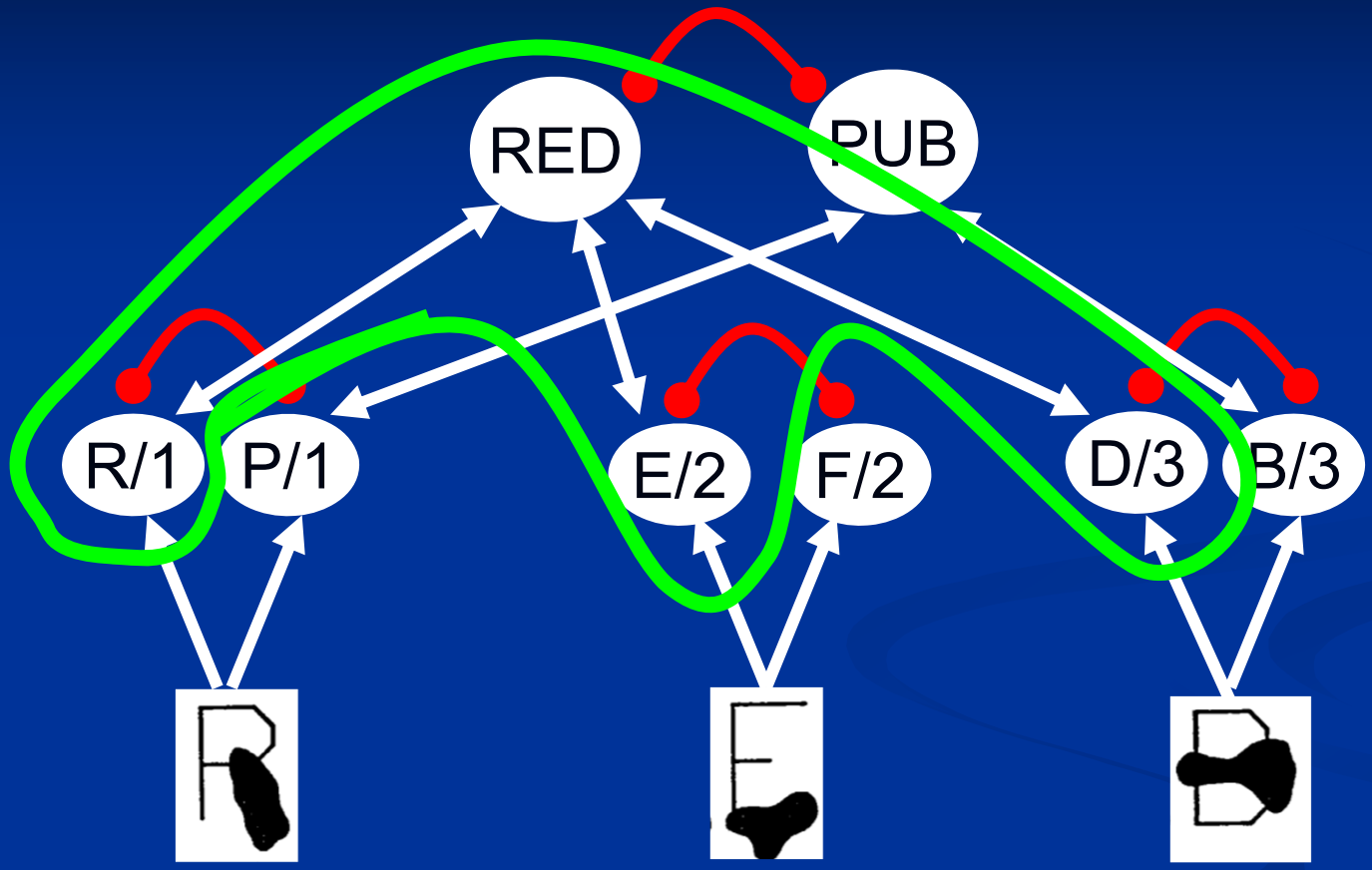


RED vs. PFB



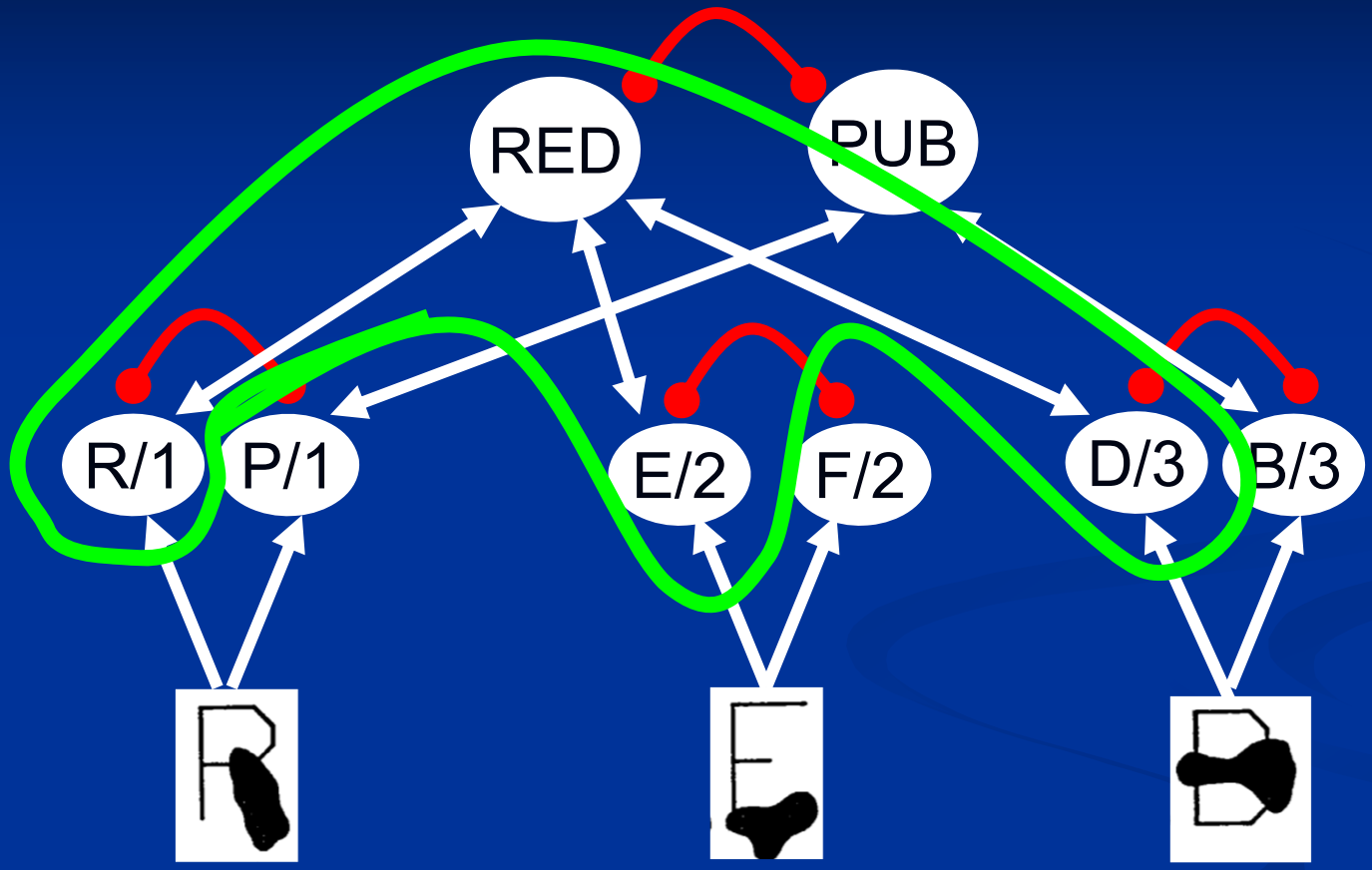
RED gets three votes, PUB only gets two

RED vs. PFB



So you get a ***stable coalition*** involving RED - and this is what reaches consciousness

RED vs. PFB



This is how *context influences perception*:

All of the letters vote for each other through their shared word node

Example of data predicted

- This model accounts for a lot of data, and predicted new data no one would have thought of measuring without the model: ***SLNT***
- SLNT is a *non-pronounceable non-word,* but it has a lot of friends at the word level
- The model predicts that there should be a superiority effect for SLNT.
- They tested this in UCSD Psychology sophomores and got the predicted effect

So, how does the brain make the mind?

- Through the concerted action of billions of nerve cells, that work together to interpret the world.
- When they activate, compete, and then settle into a stable coalition, that's what we perceive
- But there is a lot going on "under the hood" that we never experience - e.g., the temporary activation of PUB.

Summary

- Cognitive Scientists are trying to understand how the brain works
- Some of us do that by building models that "do the same things people do."
- Models are cool because you can do things to them you can't do to people: like take them apart to see why they work!

END!!

Any questions?

Your first neural net: The Interactive Activation Model: A model of reading from print

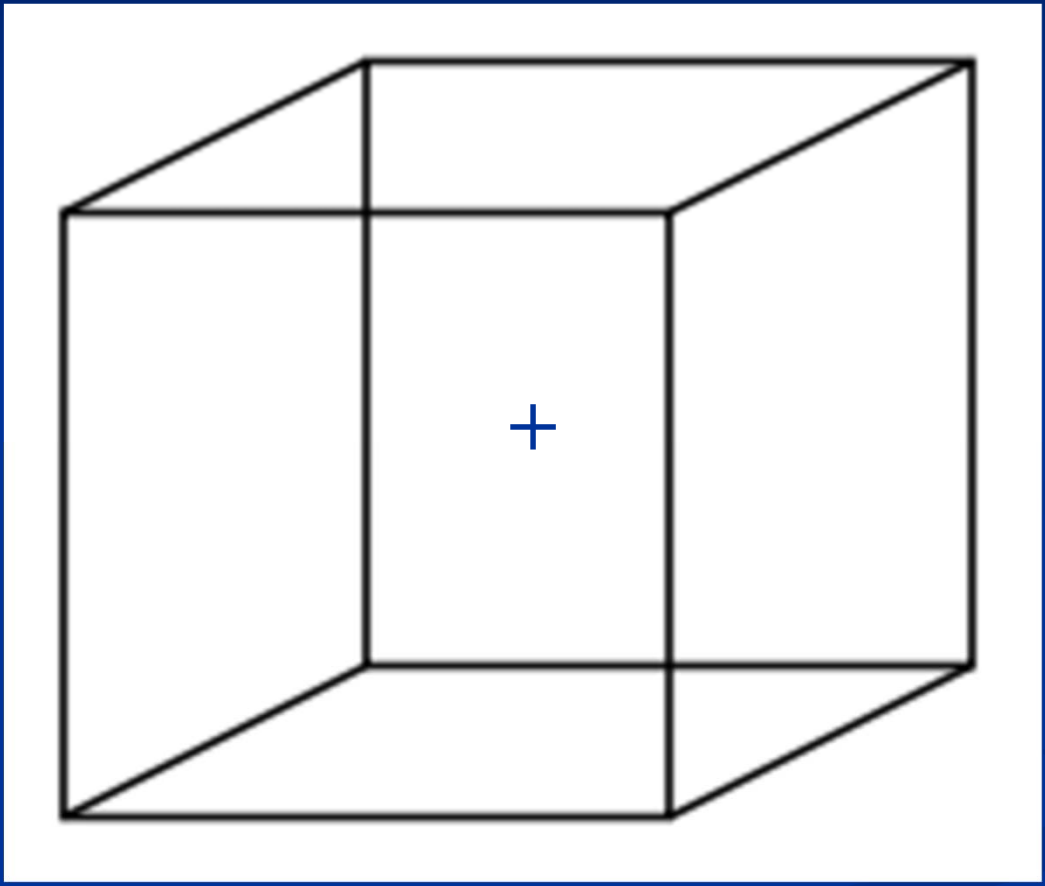
- But before we go on:
- Unbeknownst to you, *seeing things* starts with brain cells in your retina that respond to little tiny spots of light.
- Then in your visual cortex, these are put together into little bars…
- Which are put together into more complicated features…
- Until deep in your brain, you get the "Halle Berry" neuron…

Today's simulation: The Necker Cube
But first: Computer setup

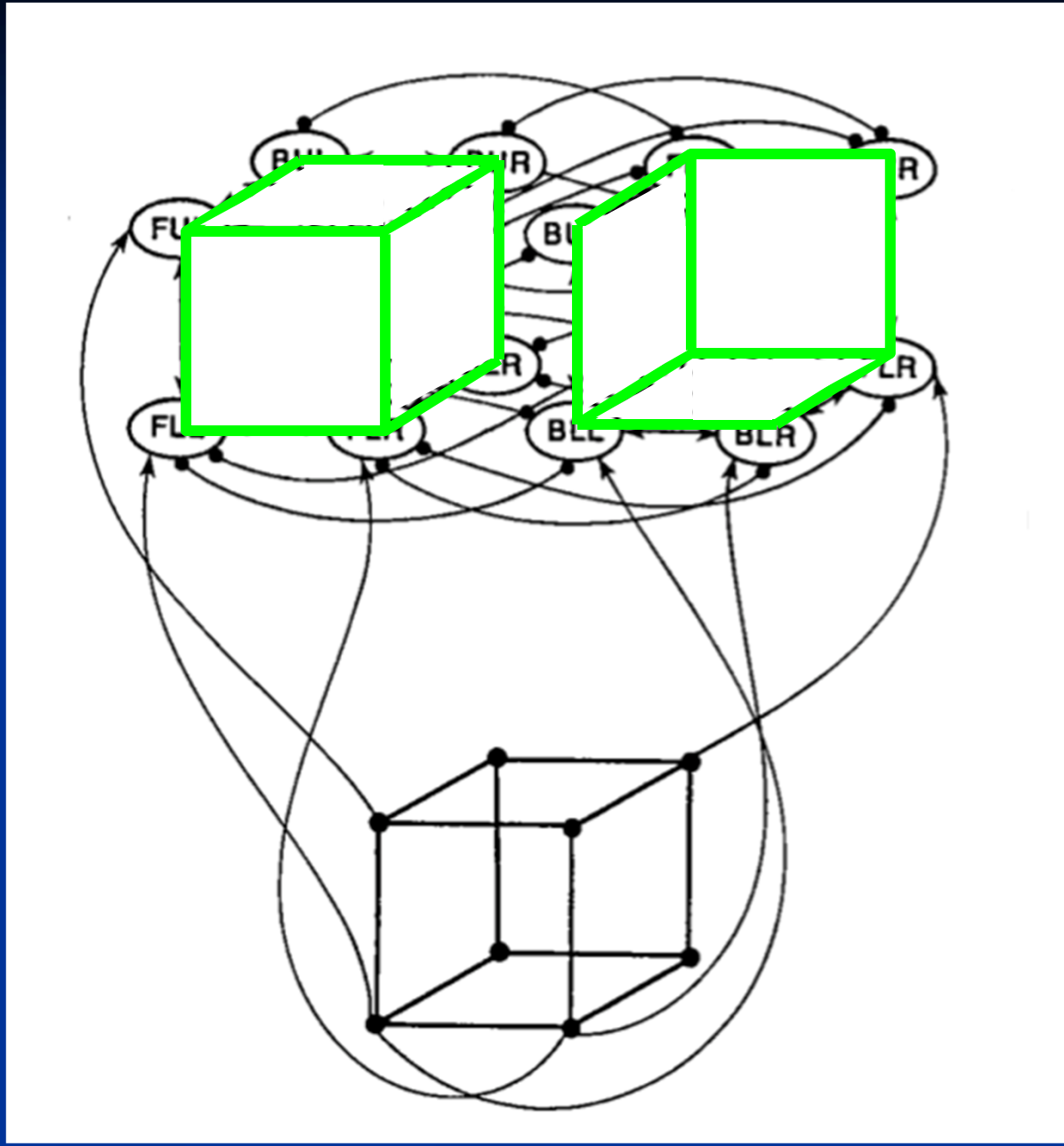
1. Apologies: According to the Academic Computing Service, you cannot keep any files on your account between logins. This is ***service???***
2. Turn over the course description I handed out and follow the instructions.

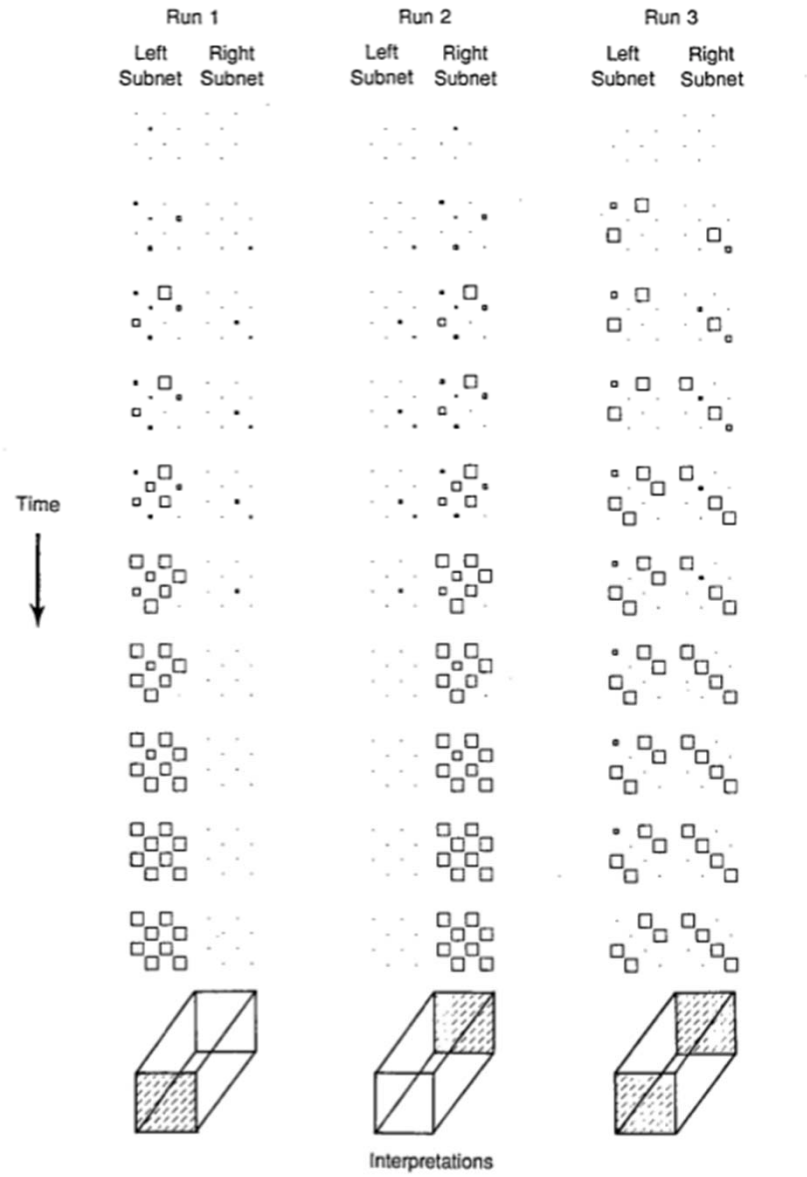
The Necker Cube:

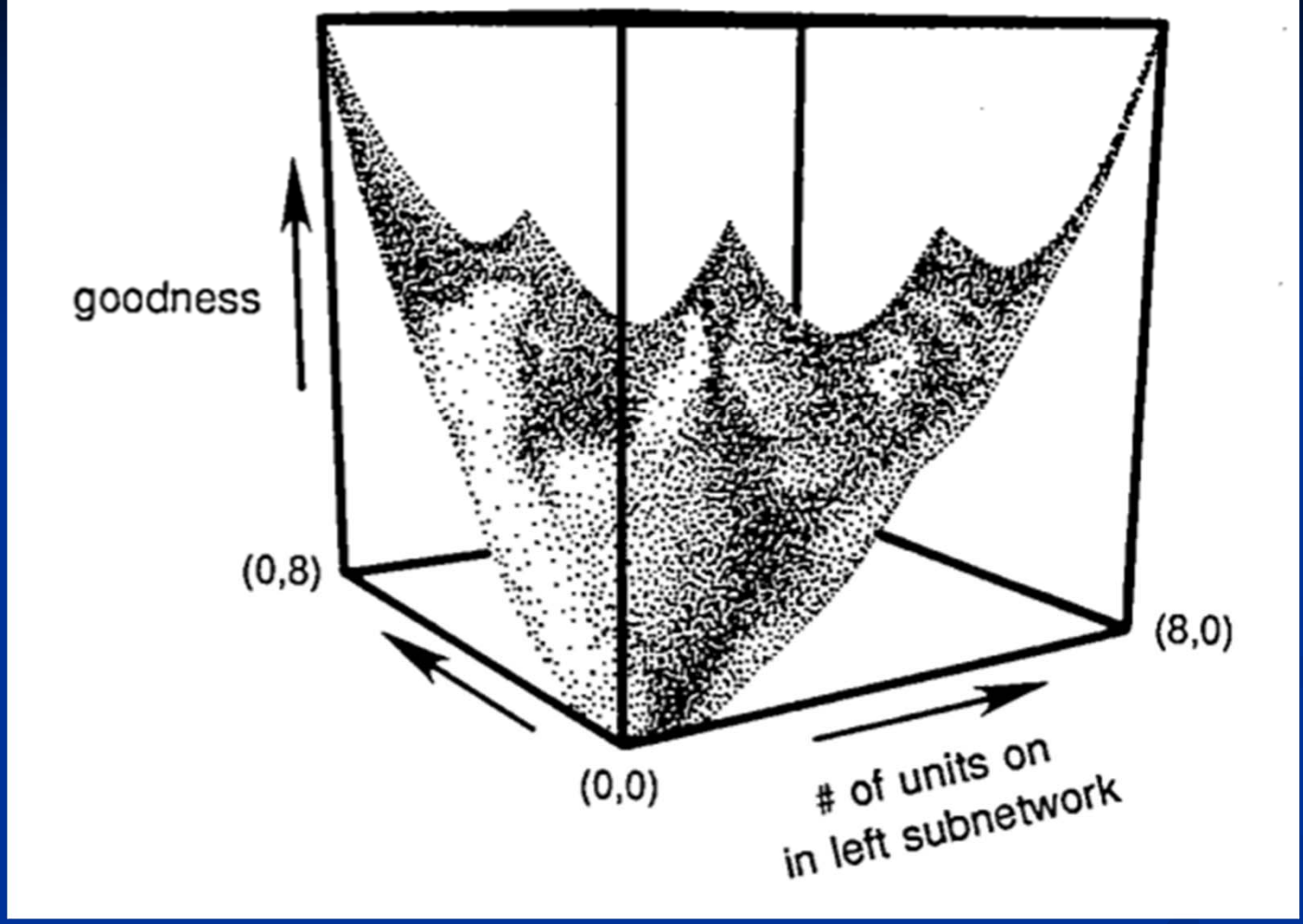
Stare at the fixation cross (+)

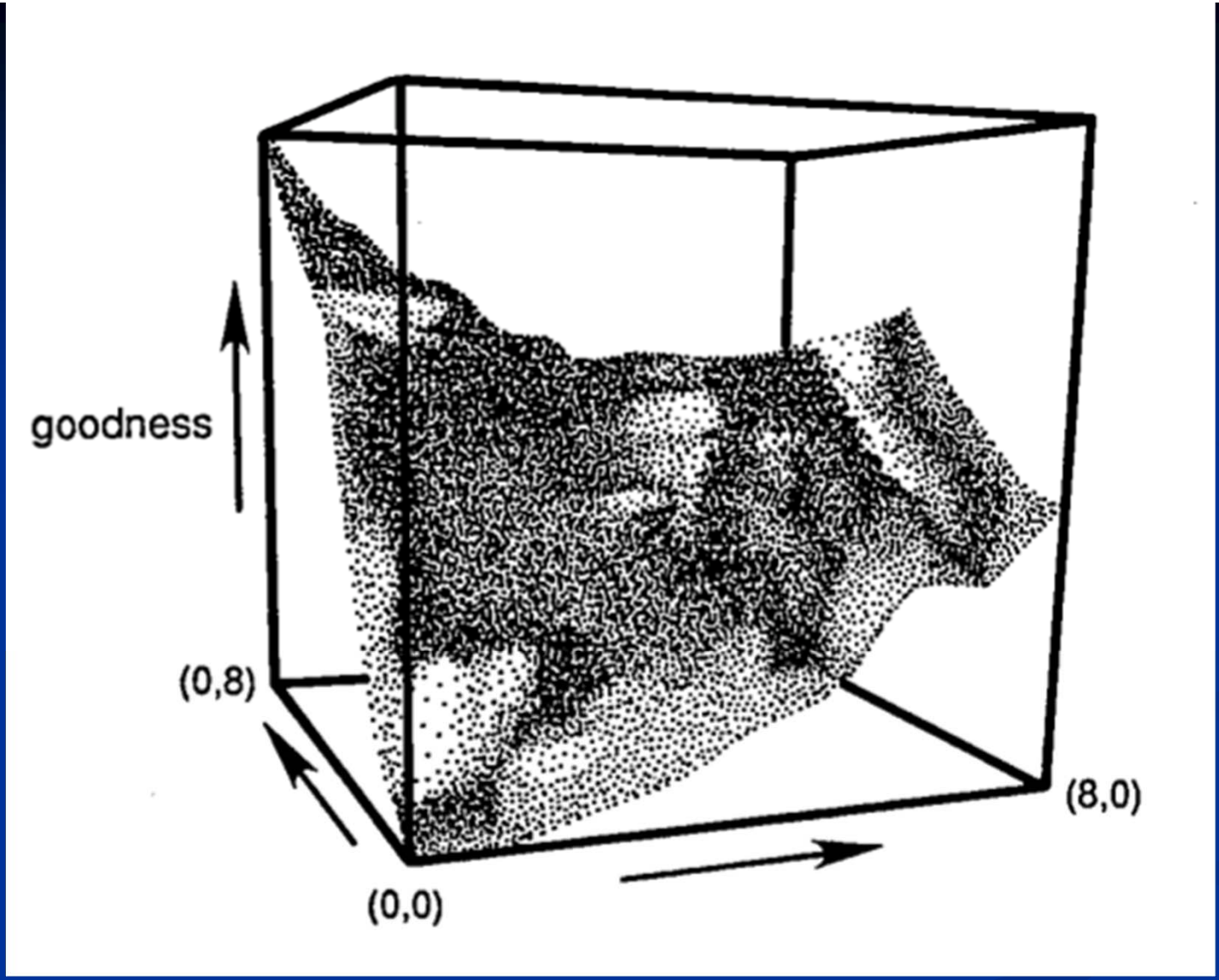


- The left cube is the down-left-facing one
- The right cube is the up-right-facing one
- Each vertex is labeled by its value – a ***hypothesis*** about that vertex.
- E.g., FLL represents the hypothesis that this particular input is the front lower left vertex.
- The units are wired up according to ***constraints*** between the hypotheses.




Run 1
Left Subnet
Right Subnet
Run 2
Left Subnet
Right Subnet
Run 3
Left Subnet
Right Subnet
Time
Interpretations


goodness
(0,8)
(8,0)
(0,0)
# of units on
in left subnetwork


goodness
(0,8)
(0,0)
(8,0)

Summary

The Necker Cube simulation showed:

1. How a network can be "hand-constructed" by:
   1. Following the "unit/value" principle
   2. Wiring up the units according to the constraints
2. How spreading activation and inhibition can "make a decision": The two networks *competed* to win via the negative links between them