Neural Representations

AI 109

Richard Kelley

Last Time

- Knowledge Representation
- Symbolic Representations
- Logical Inference

Today

- Logic Demo
- Brains, Neurons, Neural Networks
- Neural Representation

Logic Demo

Combinatorial Explosion

- If we represent facts using symbolic logic, our programs quickly become unusable – the program has to consider too many possibilities.
- Example
  - A = “The alarm is sounding”
  - B = “There is a burglary”
  - C = “The police are notified”
  - B → A = “If there is a burglary, the alarm is sounding”
  - A → P = “If the alarm is sounding, the police are notified”

Combinatorial Explosion

- What are the “possible worlds” in our model?
  1. ¬B, ¬A, ¬P (B false, A false, P false)
  2. ¬B, ¬A, P (B false, A false, P true)
  3. ¬B, A, ¬P (B false, A true, P false)
  4. ¬B, A, P (B false, A true, P true)
  5. B, ¬A, ¬P (B true, A false, P false)
  6. B, ¬A, P (B true, A false, P true)
  7. B, A, ¬P (B true, A true, P false)
  8. B, A, P (B true, A true, P true)
- Which of these are possible according to our rules?
  - World 5 isn’t possible because it doesn’t fit with B → A.

Combinatorial Explosion

- If you give "partial information" to an AI that uses symbolic logic, then it has to consider ***exponentially*** many possible worlds:
  - If the system has *n* variables, then there are 2^n possible worlds.
  - 2^n gets big *really really* fast:

n	2^n
1	2
10	1024
50	1,125,899,906,842,624
266	More than atoms in universe

- The term for this is ***combinatorial explosion*** – the number of possibilities explodes and makes it impossible for programs to run.

Logic feels like a dead end…

What should we try next?

Idea

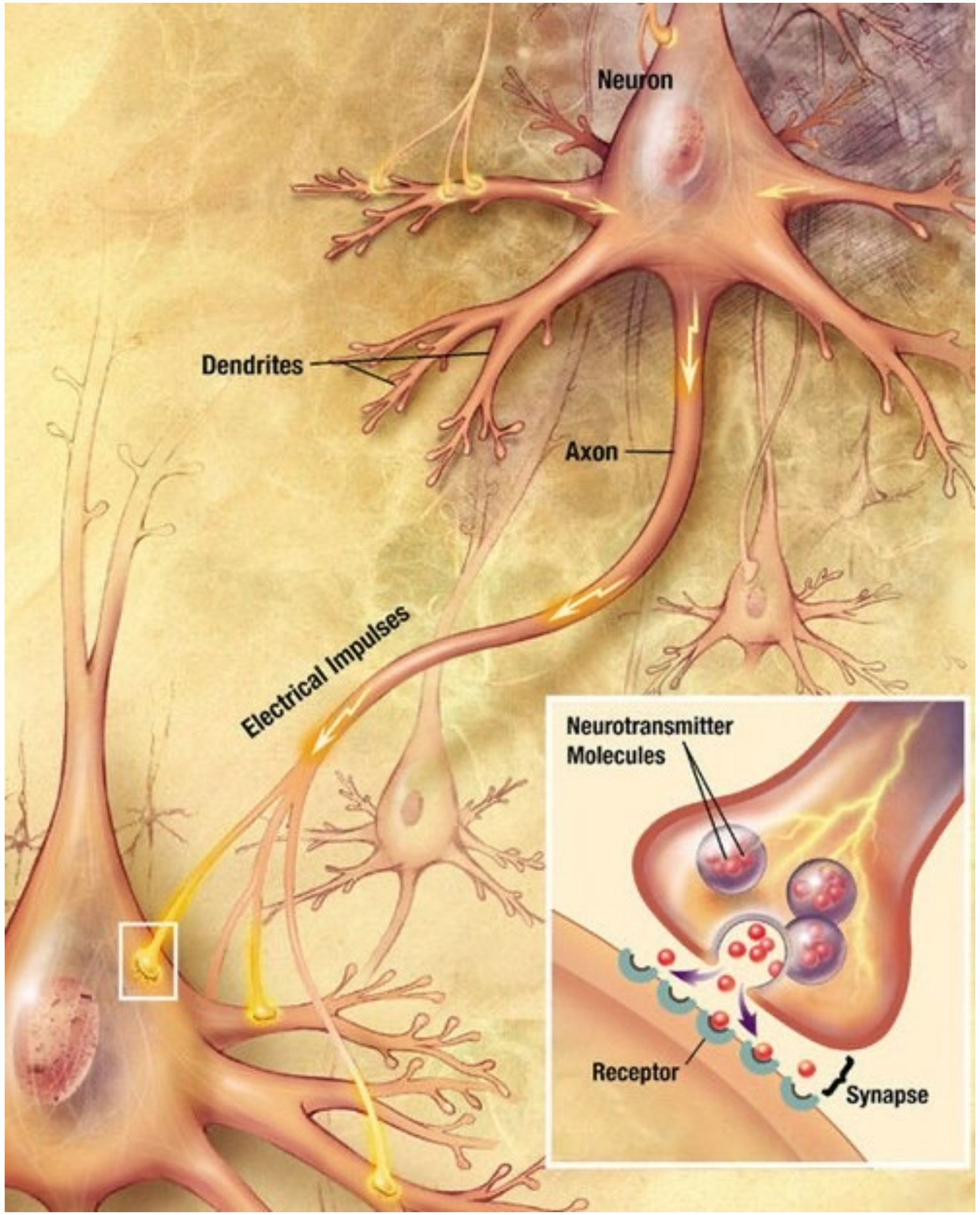
- Intelligence Seems to be Related to Brains.
- If we simulate how brains work, maybe we'll get intelligence?

How do Brains Work?

- Rough idea
  - The brain is made of cells called *neurons*.
  - Neurons send and receive electrical and chemical signals to each other.
  - This seems to have something to do with thinking.
    - If you break the brain, that frequently seems to make thinking hard/impossible.

Neurons

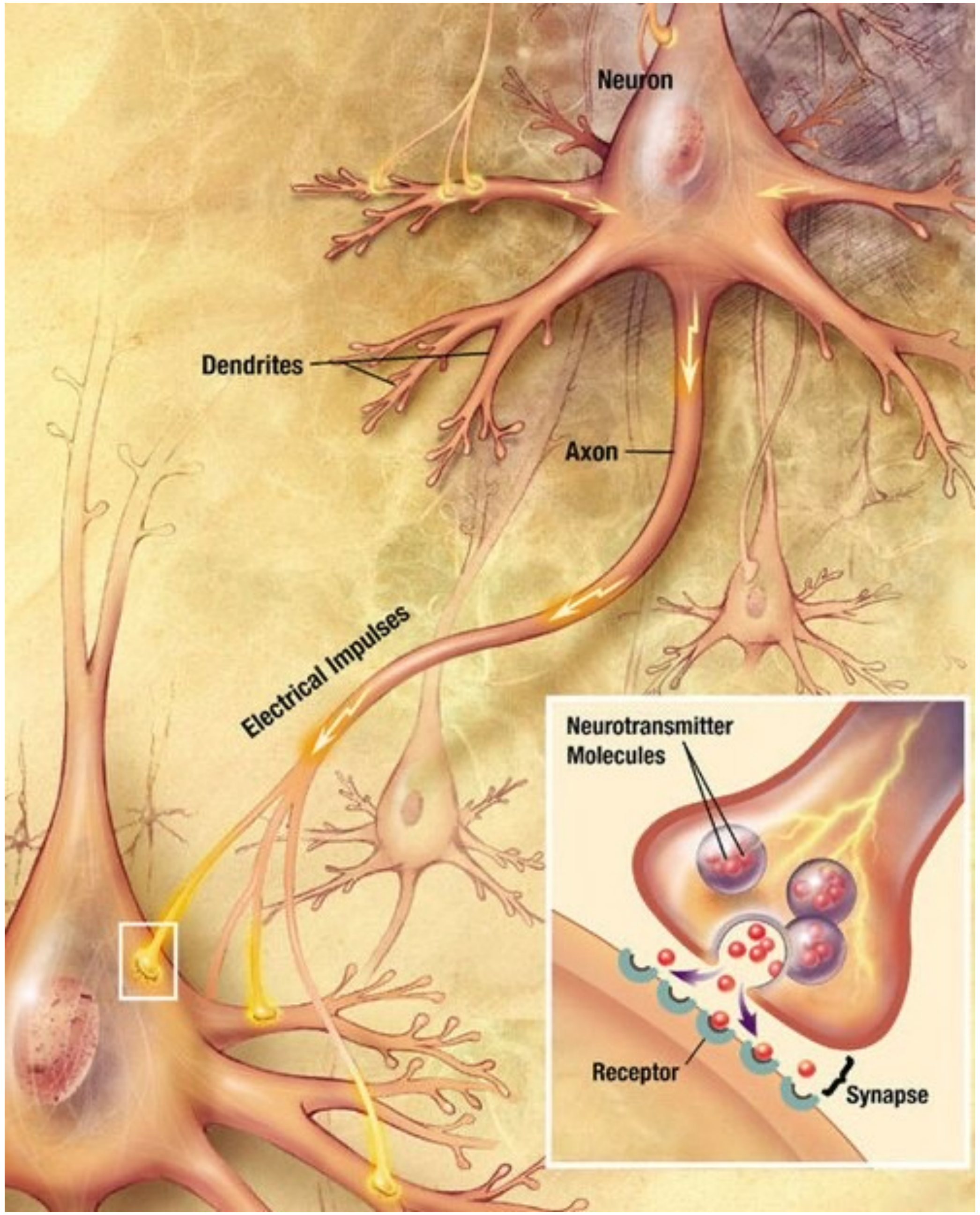
- Axons
  - Carry signals away from cell body
- Dendrites
  - Receive signals
- Impulses
  - The signals a neuron sends. These seem to “fire” in pulses.
- Thresholds
  - It seems like some threshold needs to be met before a neuron “fires” and sends signals to its neighbors.



https://en.wikipedia.org/wiki/Neuron#/media/File:Chemical_synapse_schema_cropped.jpg

Neurons

- Thresholds
  - It seems like some threshold needs to be met before a neuron “fires” and sends signals to its neighbors.
- ***Learning*** or ***adaptation*** seems to involve adjusting how the neurons are connected, and what the firing threshold is.



https://en.wikipedia.org/wiki/Neuron#/media/File:Chemical_synapse_schema_cropped.jpg

Can we model this?

- What does it mean to “create a model?”
- In this class, a “model” is always a *mathematical* model that can *run on a computer*.

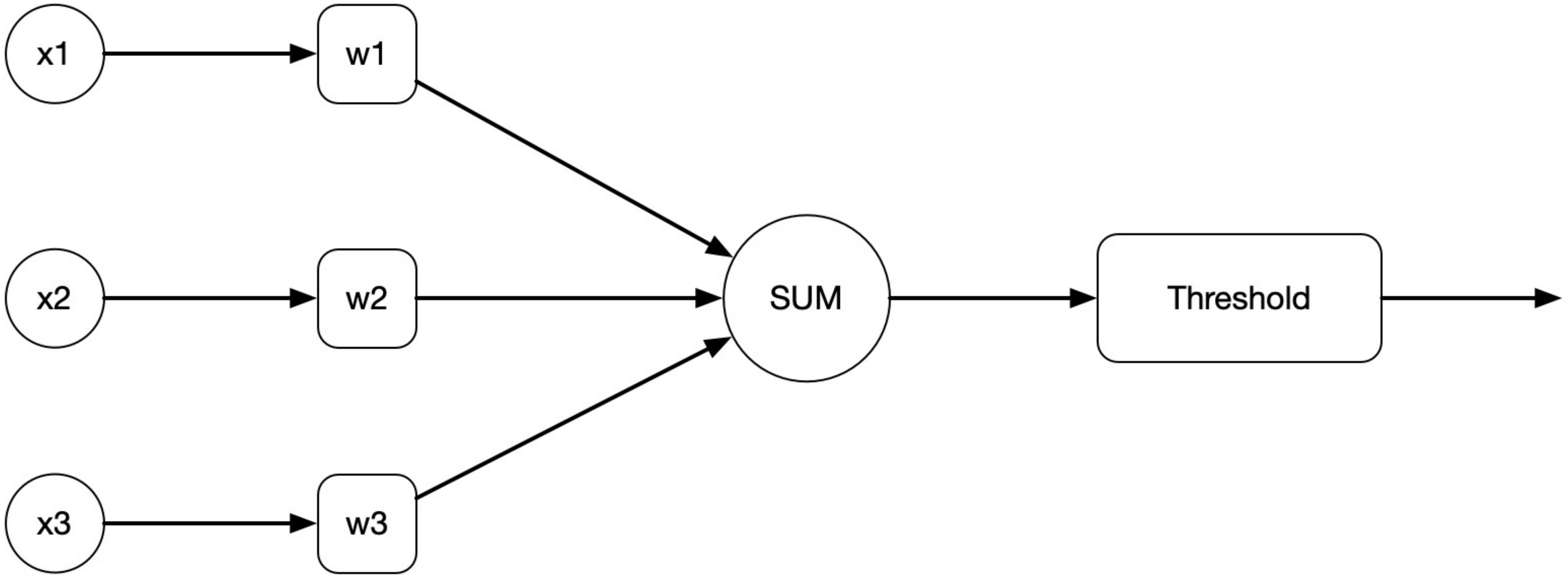
Can we model this?

- Let’s think about what it would take to build a simple model of a single neuron.
- A neuron
  - Takes some inputs,
  - Combines them (how is it combined?)
  - Compares the combined result against a threshold,
  - if the result is bigger than the threshold, then spike.

(Spiking) Neural Net Demo

A Simpler Model

- The spiking model is too complicated (and wouldn’t have worked on a 1950s computer.
- There’s a simpler alternative, called the ***perceptron***:



Rosenblatt's Perceptron

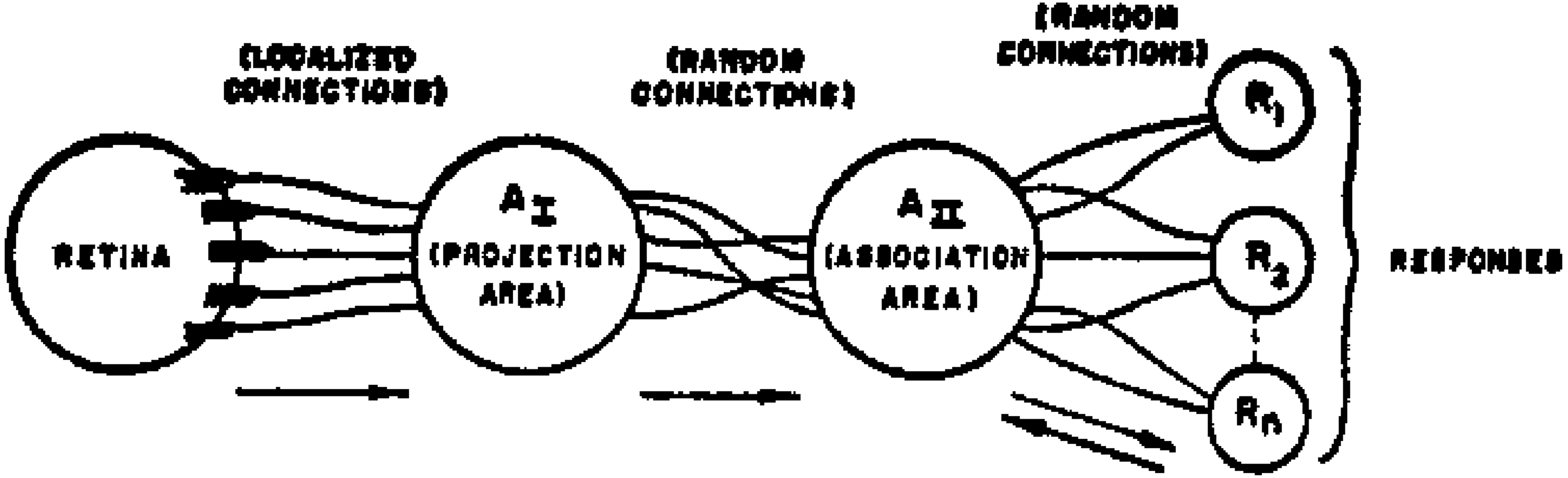


FIG. 1. Organization of a perceptron.

Rosenblatt, F. (1958). *The perceptron: A probabilistic model for information storage and organization in the brain*. **Psychological Review, 65**(6), 386–408.

The New York Times

Electronic 'Brain' Teaches Itself

July 13, 1958

What can we do with this?

- **Classification**
  - Recognize Images. The network outputs a 1 when the input takes a certain shape.
  - Each pixel in the image is an input.
  - Requires that the system "learn" to distinguish one thing from another.
    - Adjusting weights and thresholds.