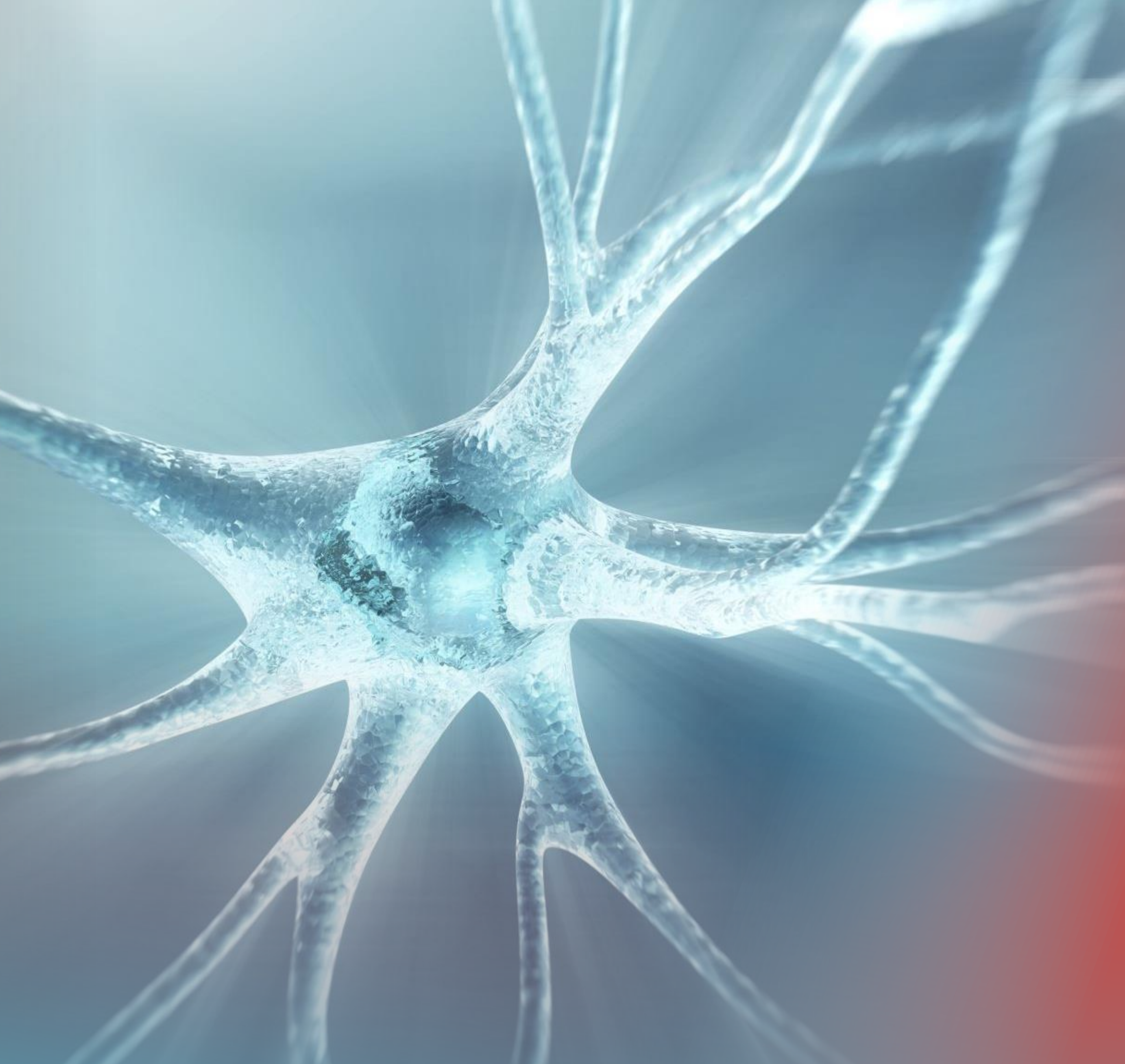
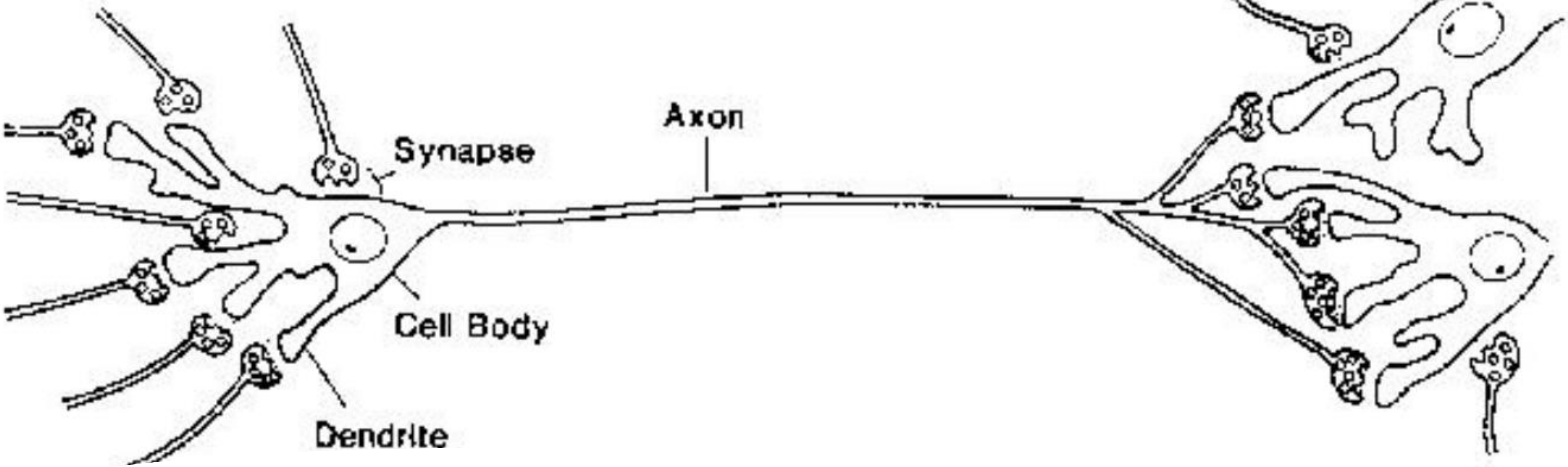
Neural Networks

Historical Foundations and Neurobiological Inspirations

CSCI 4850/5850





What’s so special about neurons?

- Slow (milliseconds), low bandwidth, and imprecise (noise)
- Unreliable (cell death), but networks degrade gracefully
- Massively parallel – 1e11 neurons, 1e14 synapses
- Continuous operation over time
- Distributed Computation
- Connection pathways topologically structured
- Learning via synapse modification

Artificial Neural Networks and Artificial Intelligence

- The computer/brain metaphor has a long history in artificial intelligence

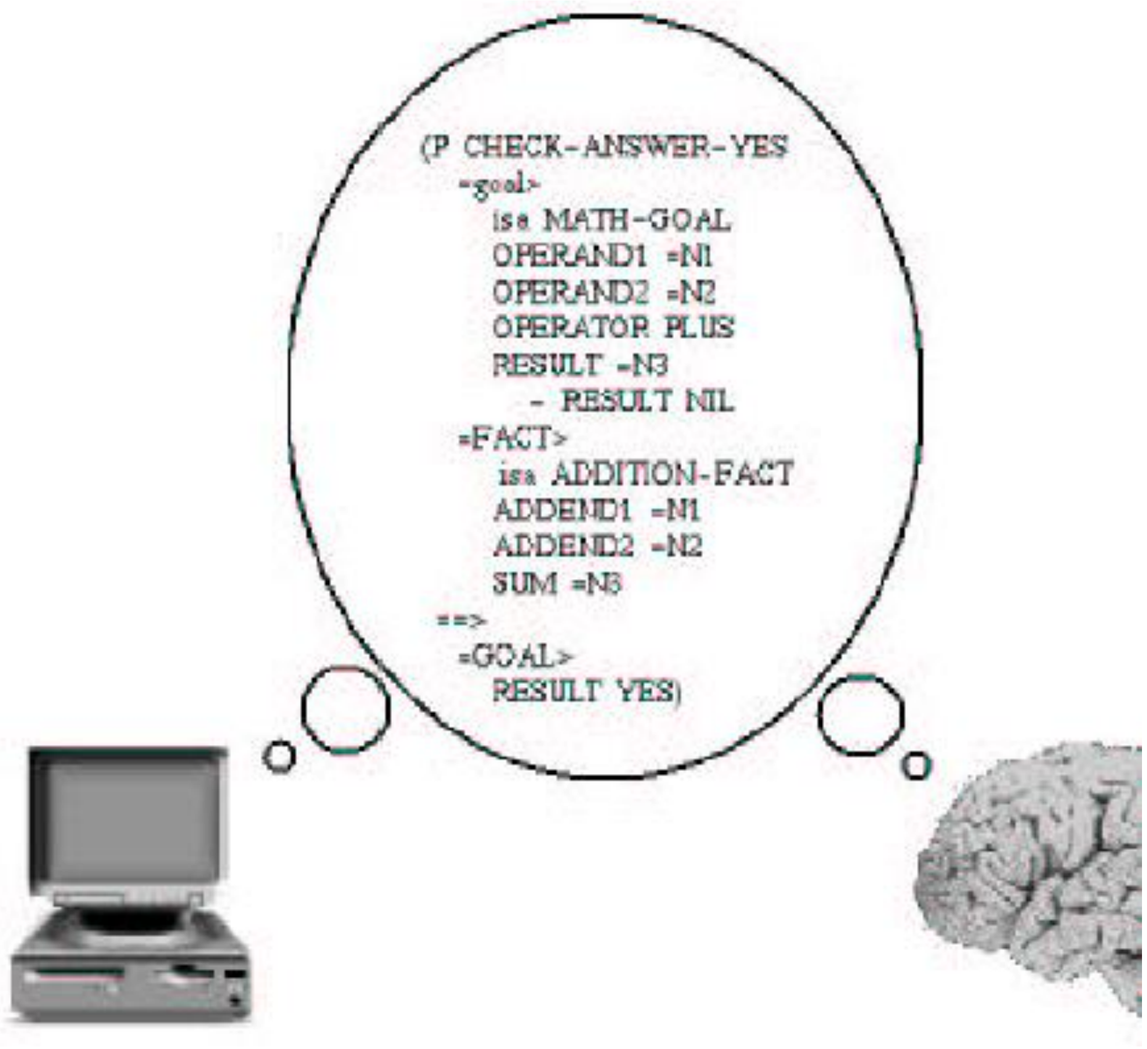
- What can a computer do that a human cannot?

- What can a human do that a computer cannot?

- Is one smarter than the other?

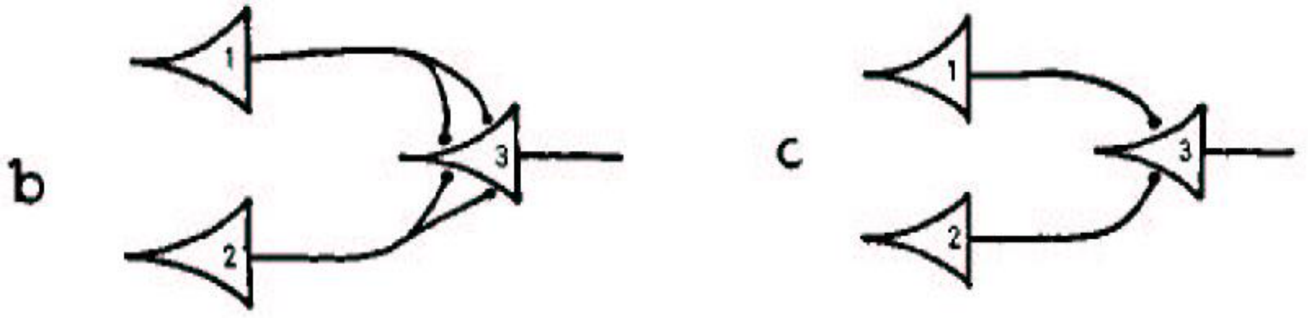
- Is the difference only software?

- Does the architecture make a difference in what the computer/brain can do?



A very brief history of neural computation…

“Since the beginning of the field of computer science, the computer has been seen as a kind of brain, and the brain has been seen as a kind of computer” (McCulloch and Pitts, 1943)



The main question was *how* the brain computed: how does it give rise to complex behaviors?

What kind of computation?

100-step constraint (Feldman and Ballard, 1982)

- Given the speed of individual neurons…
  - Hundreds of pulses per second…
  - Basic mental events take less than a second for humans to perform (eg. recognize your mother)…
  - This would need to be performed in less than 100-steps if the computation was serial!

Do as much in parallel as possible

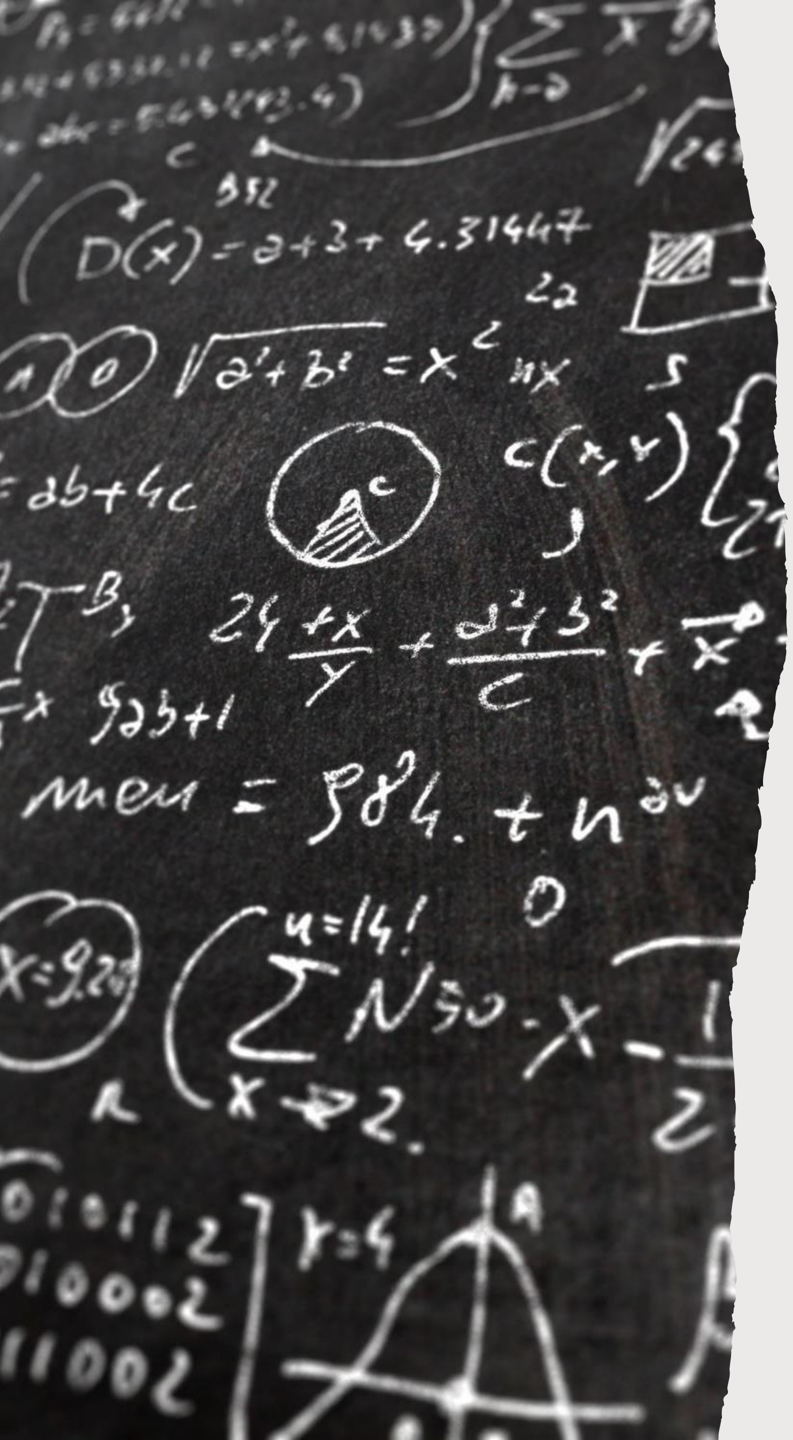
Make use of information as soon as it's available

Distribute the task across many redundant processing elements

Focus on learning from experience

The early historical context…

- Turing – 1936 – "On Computable Numbers"
- Mauchly and Eckert – 1943-46 – ENIAC Project
- Von Neumann – 1945 – "First Draft of a Report on the EDVAC"
  - Stored-program architecture
- 1952 – UNIVAC predicted presidential election results
- Neuroscience inspires new ideas in communications and control theory
  - Cybernetics – or Control and Communication in The Animal and The Machine (Wiener, 1948)
  - The Organization of Behavior (Hebb, 1949)
  - In Search of the Engram (Lashley, 1950)

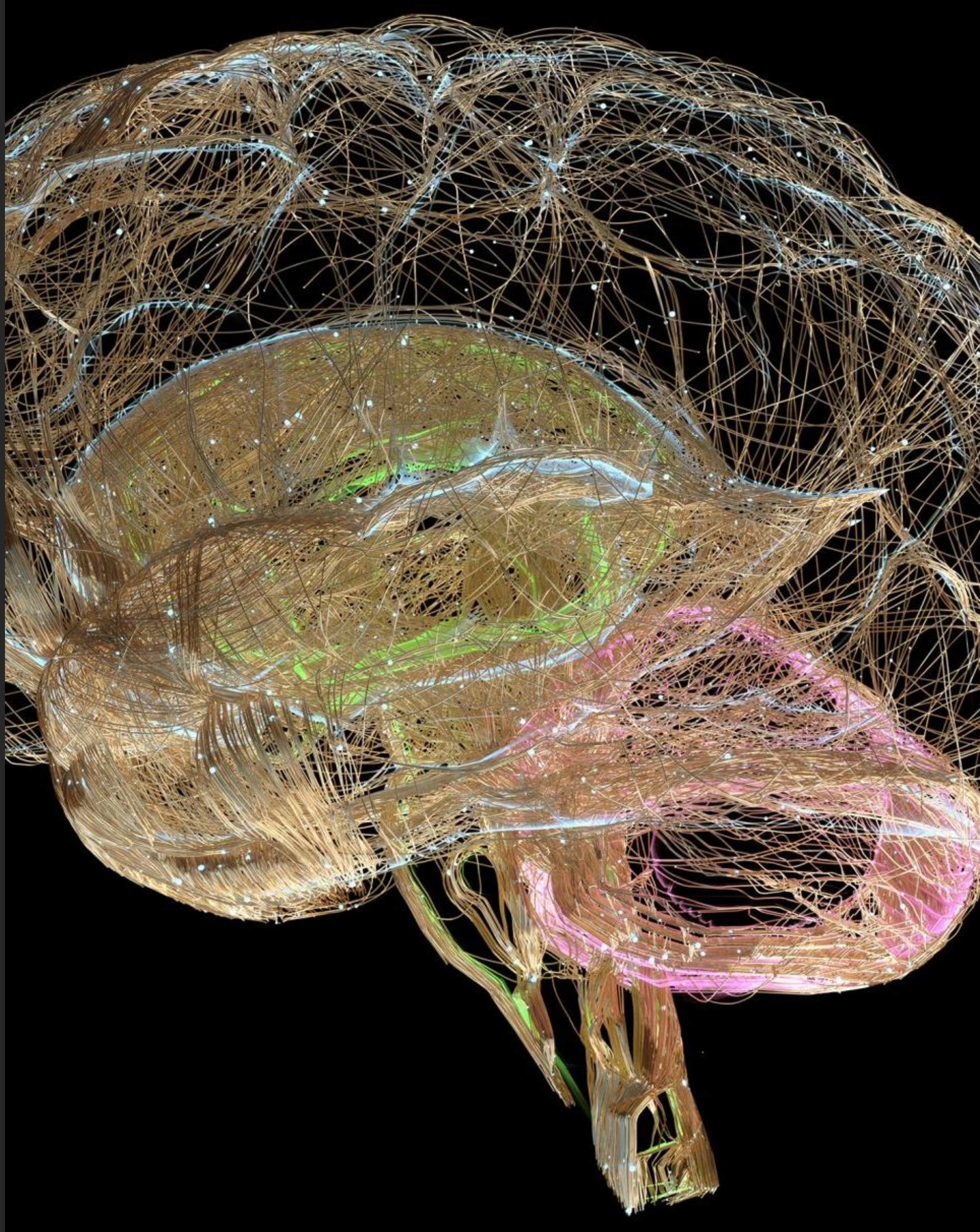


Brains –versus- Computers

“Thus the nervous system appears to be using radically different systems of notation from the ones we are familiar with in ordinary arithmetics and mathematics: instead of the precise systems of markers where the position — and presence or absence — of every marker counts decisively in determining the meaning of the message, we have here a system of notations in which the meaning is conveyed by the statistical properties of the message.” John von Neumann, 1958

Initial Discoveries

- Theory of Neural-Analog Reinforcement Systems and Its Application to the Brain-Model Problem (Minsky, 1954)
- Pandemonium (Selfridge, 1958)
- ADALINE (Widrow and Hoff, 1960)
  - Delta Rule
  - Memistor Corp.
- Principles of Neurodynamics (Rosenblatt, 1962)



The Dark Ages

- Symbolic systems become preferred…
- Perceptrons (Minsky and Papert, 1969)

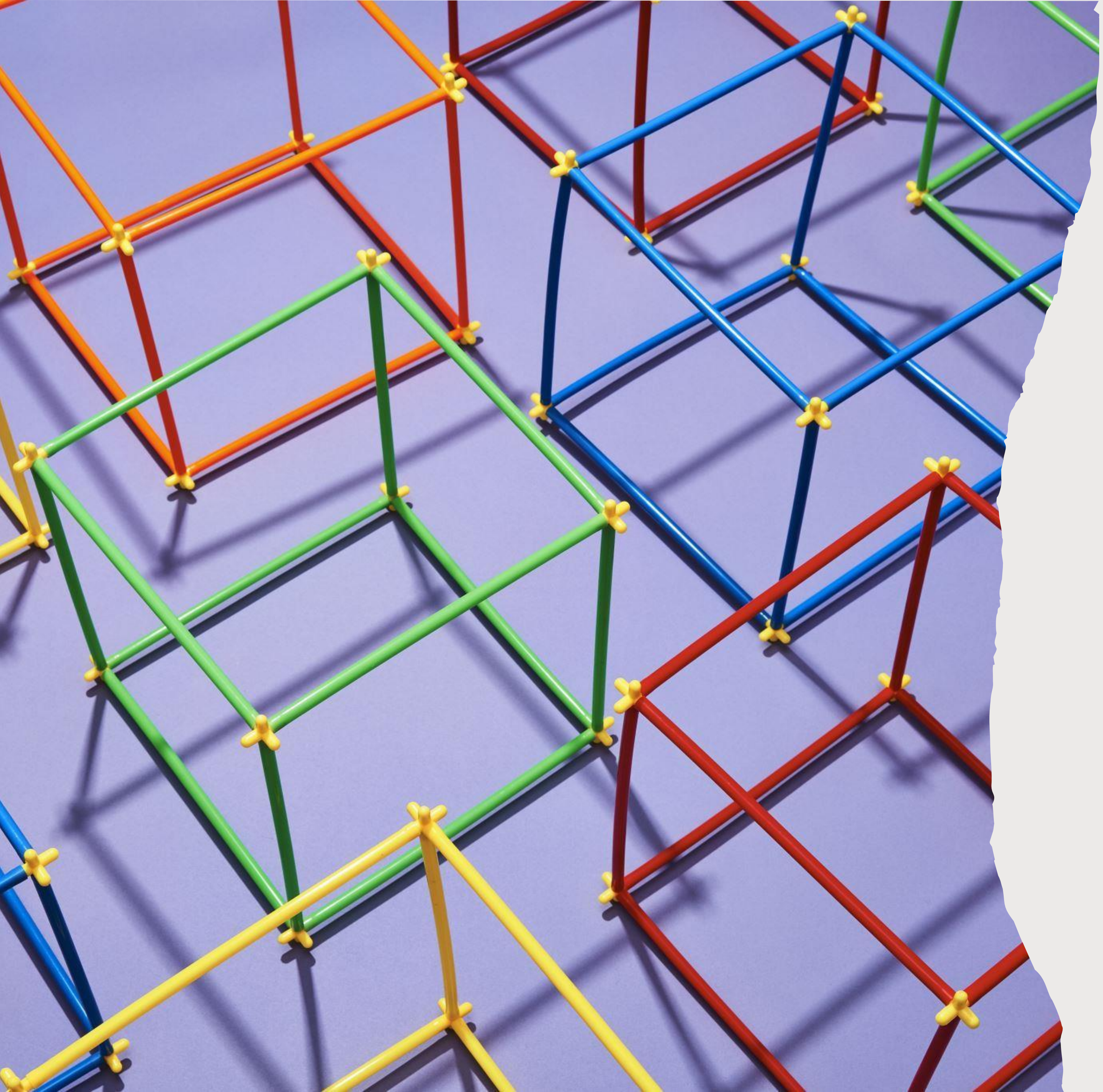
  “Perceptrons have been widely publicized as “pattern recognition” or “learning” machines and as such have been discussed in a large number of books, journal articles, and voluminous “reports.” Most of this writing (some exceptions are mentioned in our bibliography) is without scientific value and we will not usually refer by name to the works we criticize.”
- Some valid criticism, but also speculative attacks…

Passing the torch…

- Shun-ichi Amari
- James Anderson
- Michael A. Arbib
- Dana Ballard
- Andrew Barto
- Jack Cowan
- Jerome Feldman
- Kunihiko Fukushima
- Stephen Grossberg
- Douglas Hofstadter
- Harry Klopf
- Teuvo Kohonen
- Christoph von der Malsburg
- Richard Sutton
- David Willshaw

“Connectionism” Renaissance

- Neural networks and physical systems with emergent collective computational abilities (Hopfield, 1982)
  - Hopfield Network Convergence Theorem
- Parallel Models of Associative Memory (Hinton and Anderson, 1981)
  - Collected papers from a ‘79 conference and updated again in ‘89
- Parallel Distributed Processing (Rumelhart and McClelland, 1986)
  - Backpropagation Learning Algorithm



New Institutions

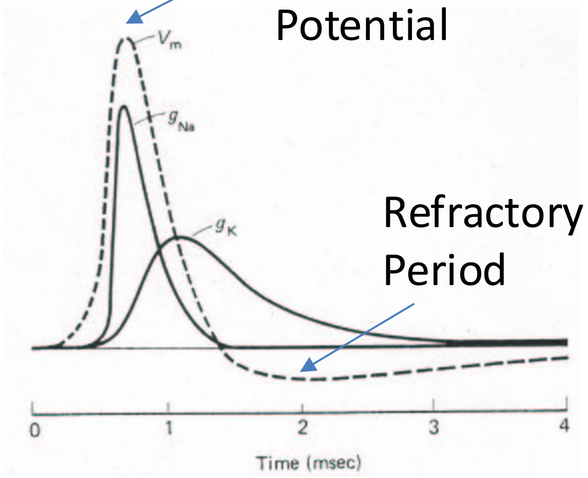
- Conferences
  - IEEE International Conference on Neural Networks (1987)
  - Neural Information Processing Systems (1987)
  - Computational Neuroscience (1992)
- Journals
  - Neural Networks (1988)
  - Neural Computation (1989)
  - IEEE Transactions on Neural Networks (1990)

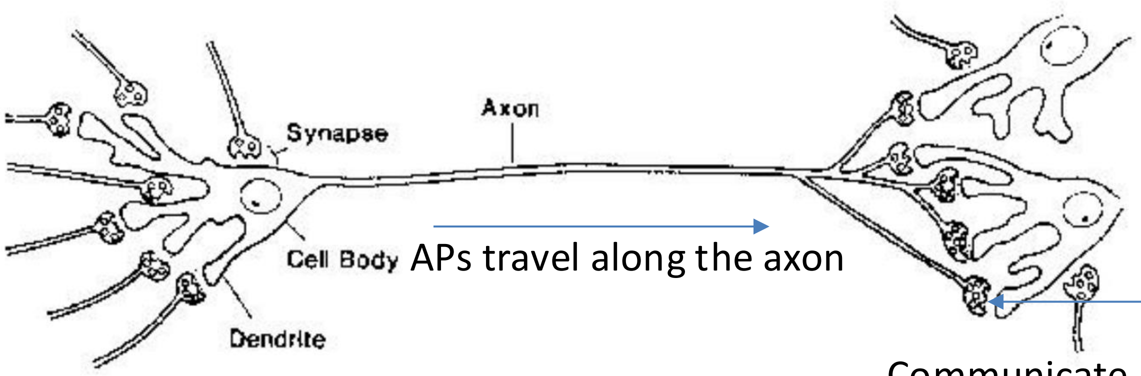
Into the Era of Deep Learning

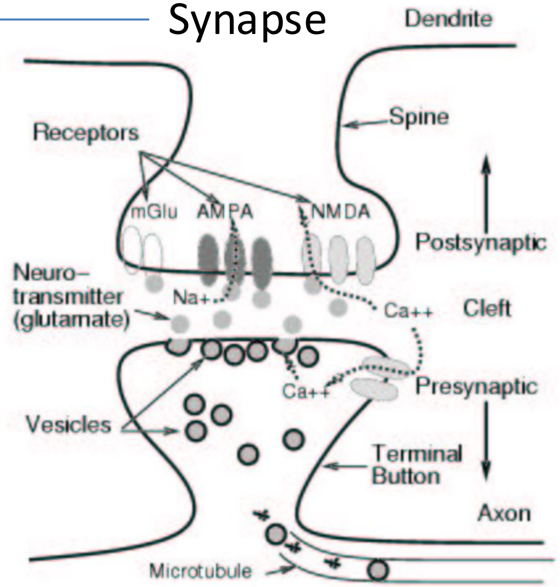
- A few key observations from the late 90’s and early 00’s
  - Weakening gradient problem
  - Minimal redundancy without sparsity
  - Second-order approximation needed
  - Numerical tricks needed for stability
- New activation functions
  - Non-weakening gradient
  - Better match to rate code for real neurons!
- New versions of backpropagation
  - Better use of 2nd order approximations
  - Adaptive and less sensitive parameters
- GPU/CPU acceleration makes big network training feasible – Gaming->Science
- Software Engineering – computational graphs - LeCun et al., 1998
- “Explosion” (relatively speaking) in application-domain success stories in early 10’s
- Now called “Deep-Learning”, the field continues to make progress today…

Biological Inspirations

- Neurons communicate:
  - *electrically* via **action potentials**
  - *chemically* via **neurotransmitters**



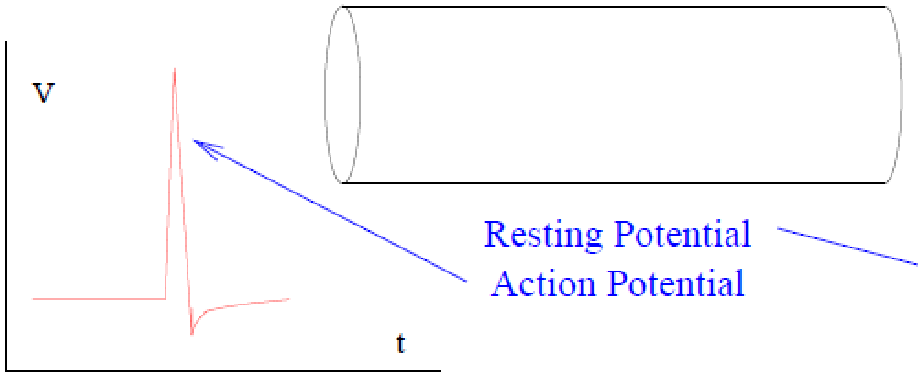


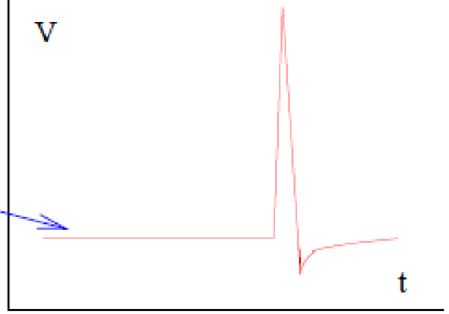


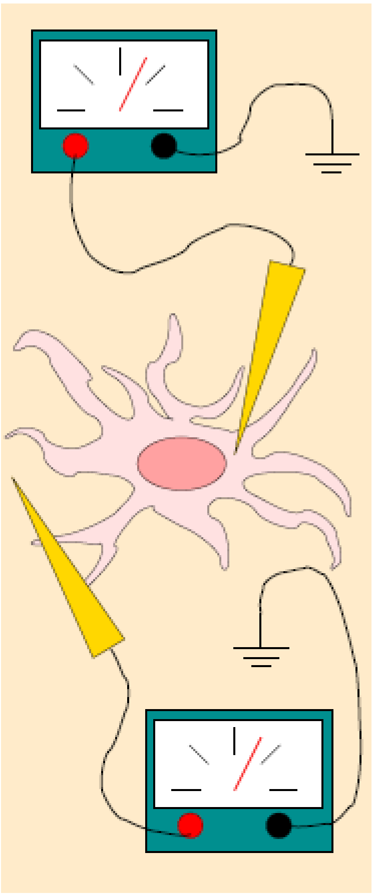
Ion Gradients and Channels

- Electrical signals from ions
  - Na^+, Cl^-, K^+, Ca^{2+}
- Ion Channels
  - Gated and Ungated
  - *Sodium-Potassium Pump*

Excitation is often called *depolarization*



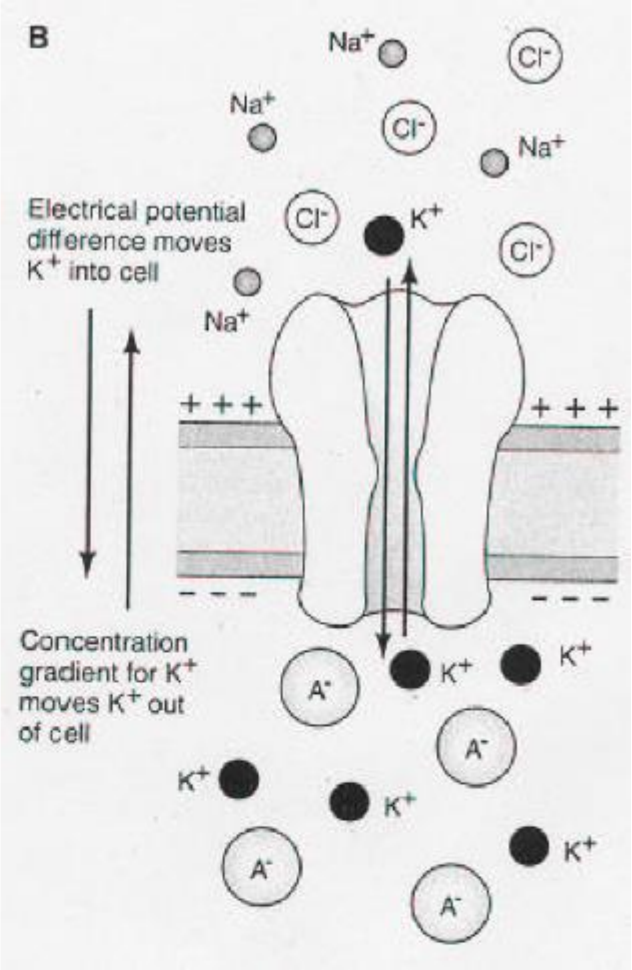




Membrane potential = $\text{Voltage}_{\text{inside}}$ - $\text{Voltage}_{\text{outside}}$

Passive Ion Channels

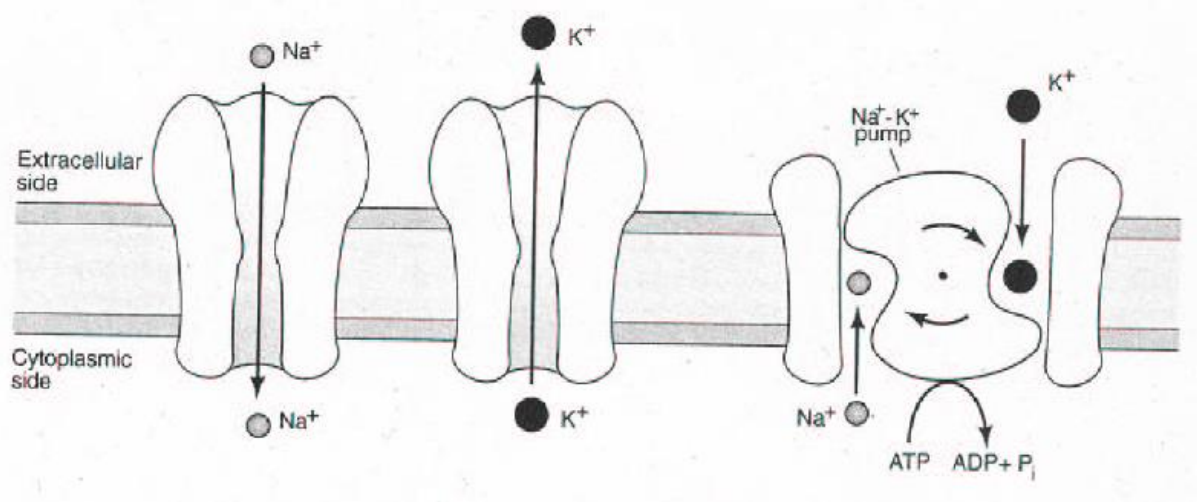
- Ions flow through passive channels driven by two (sometimes competing) forces
  - Electrostatic force
  - Diffusion
- This “competition” reaches an equilibrium point called the *equilibrium potential*
- Can be calculated for each channel using the *Nernst Equation* (below)



$$E_K = \frac{RT}{ZF} \ln \frac{[K^+]_o}{[K^+]_i}$$

Channel Types

- E_K ~= -75 mV
- E_{Na} ~= 55mV
- E_{Cl} ~= -60mV



- Lots of K inside the cell, and lots of Cl outside of the cell. An *active* pump helps keep ion concentrations constant.

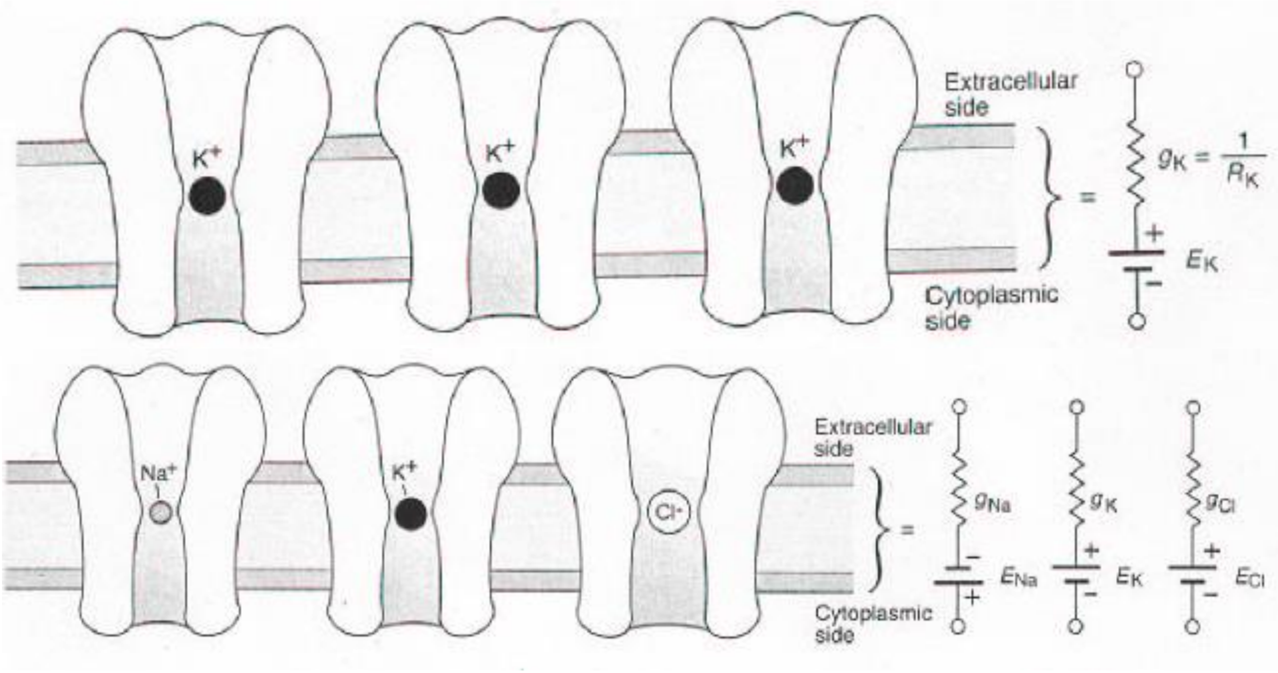
The Resting Potential

- There are many more potassium channels in a typical neural membrane than sodium channels, so the potassium equilibrium potential dominates.
- This gives rise to a *resting potential* of around -60 mV to -70 mV, and can be computed using the *Goldman Equation* (below).

$$V_m = \frac{RT}{F} \ln \frac{P_K[K^+]_o + P_{Na}[Na^+]_o + P_{Cl}[Cl^-]_i}{P_K[K^+]_i + P_{Na}[Na^+]_i + P_{Cl}[Cl^-]_o}$$

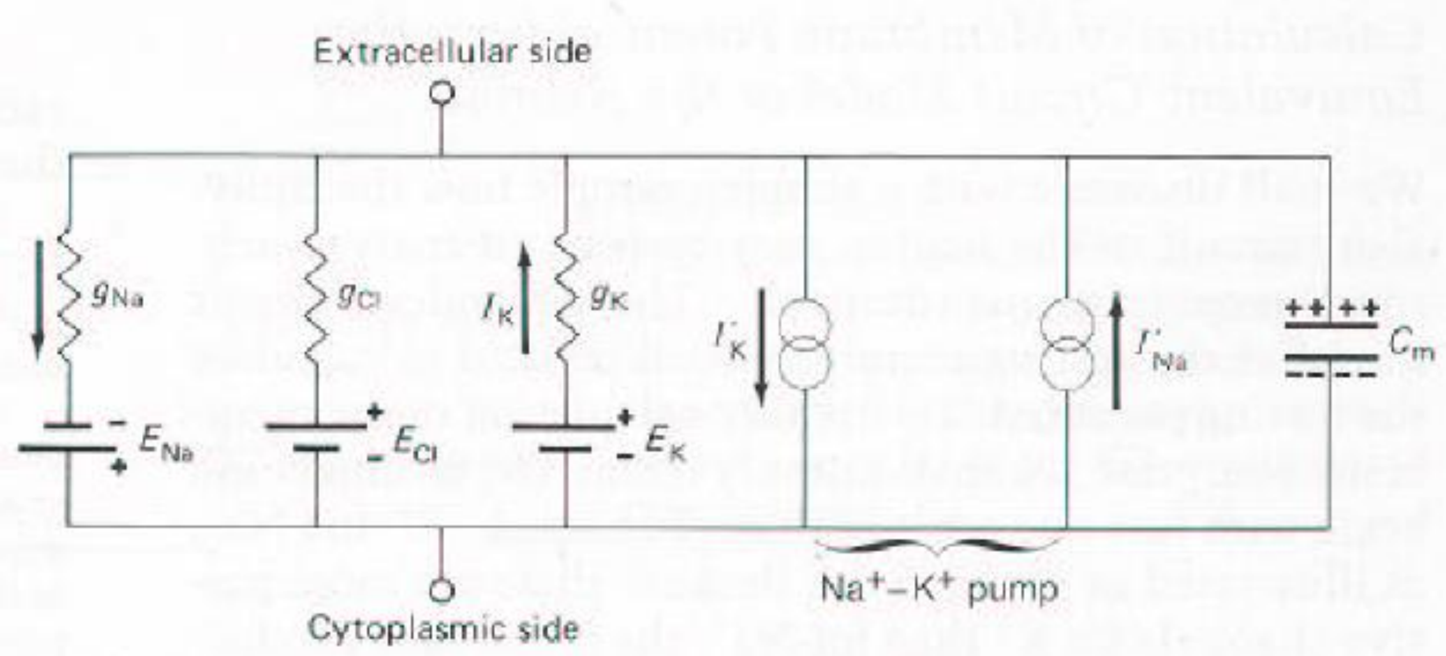
Circuit Models of Passive Channels

- The restrictive passage provided by ion channels, along with their density, limits the rate at which ions can flow across the membrane. This can be modeled by a *resistor*. Similarly, the forces that produce the equilibrium potential can be modeled by a *battery*.



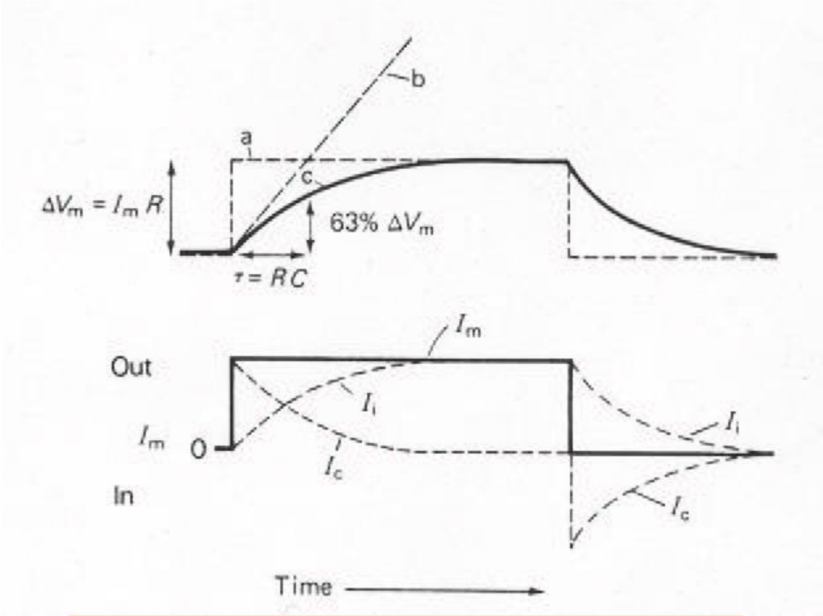
Circuit Model of the Membrane at Rest

- The bulk of the surface area of the membrane cannot be permeated by ions, introducing *capacitance* into the circuit.

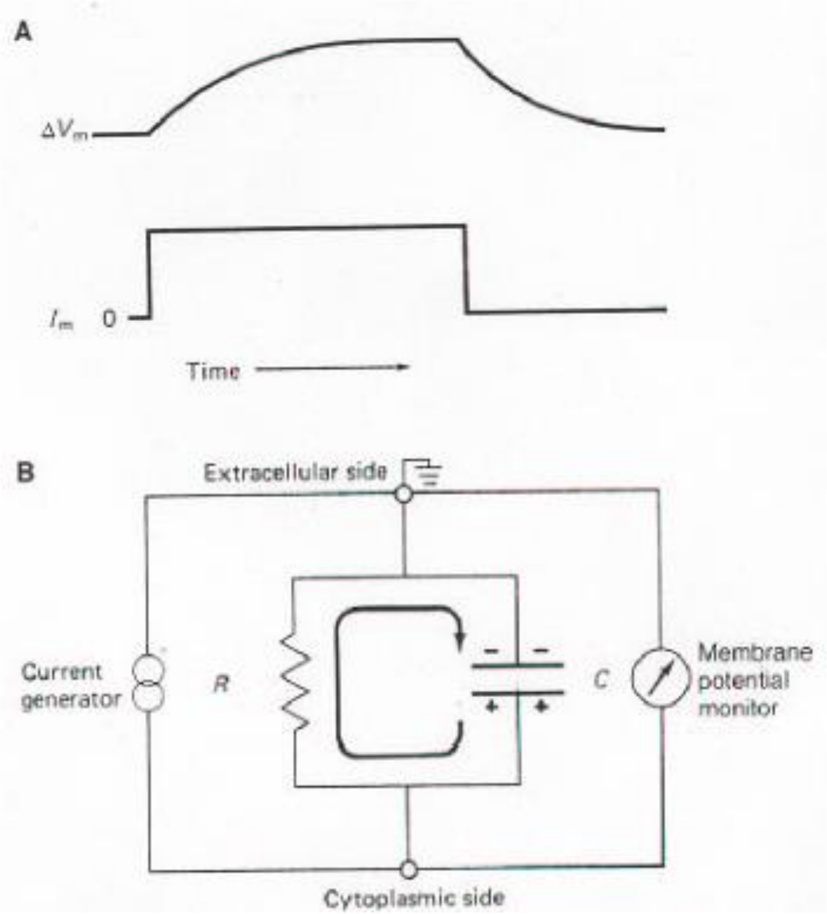


The Electrical Signal Over Time

- The capacitance in the circuit introduces a delay in the effect of injected current on the membrane potential.

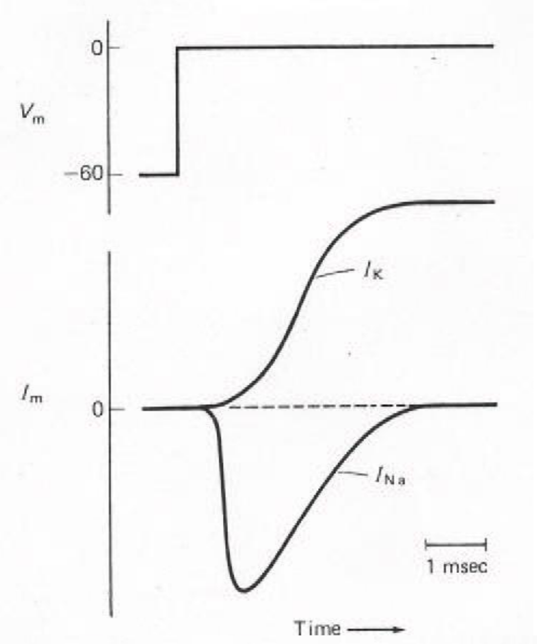


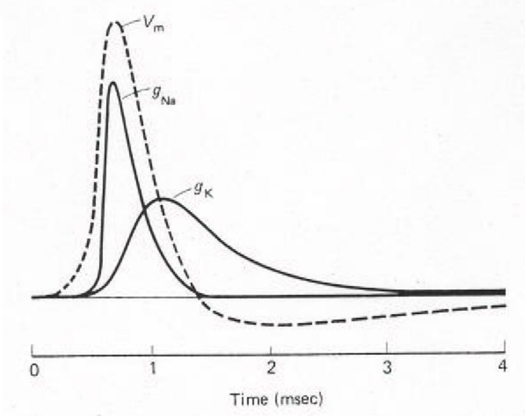
$$\Delta V_m(t) = I_m R\left(1 - e^{-t/\tau}\right)$$



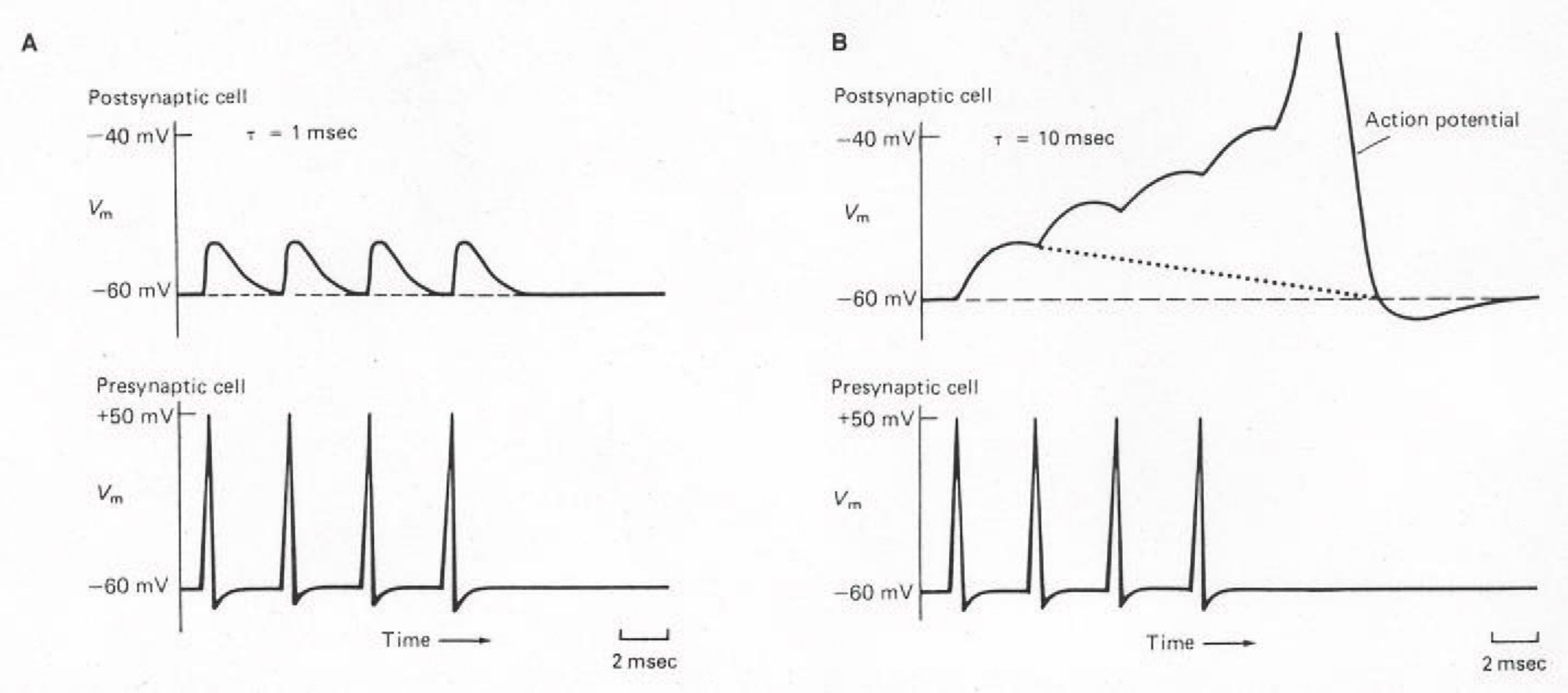
Hodgkin-Huxley Model

- Active Channels Respond to V_m
  - Sodium channels *deactivate* if the cell is depolarized for too long.
  - Potassium channels respond as well, but *more slowly*.
- The result of these active channels is an all-or-none **action potential** (or *spike*), followed by a hyperpolarizing *afterpotential*, producing a *refractory period.*
- Lab studies of channel behavior by Hodgkin and Huxley allowed them to create a model of cascading activity.





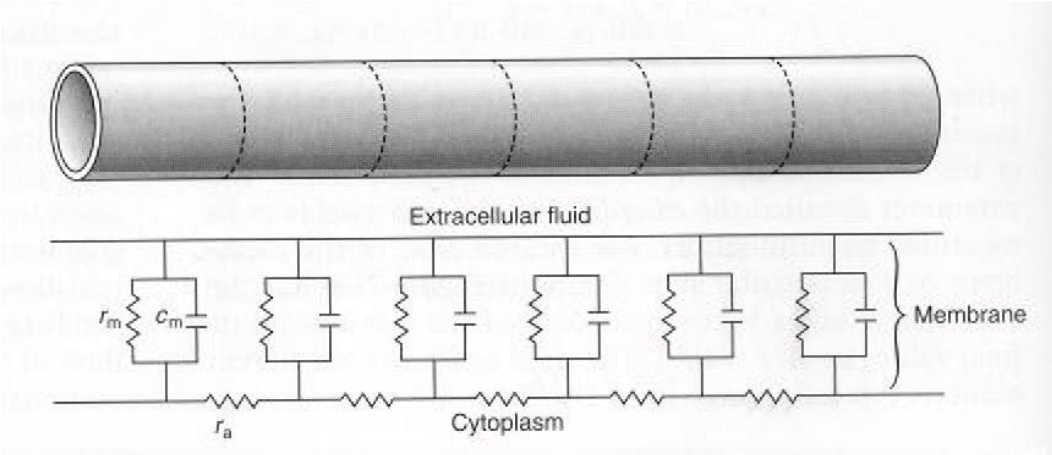
The time constant (τ) matters…

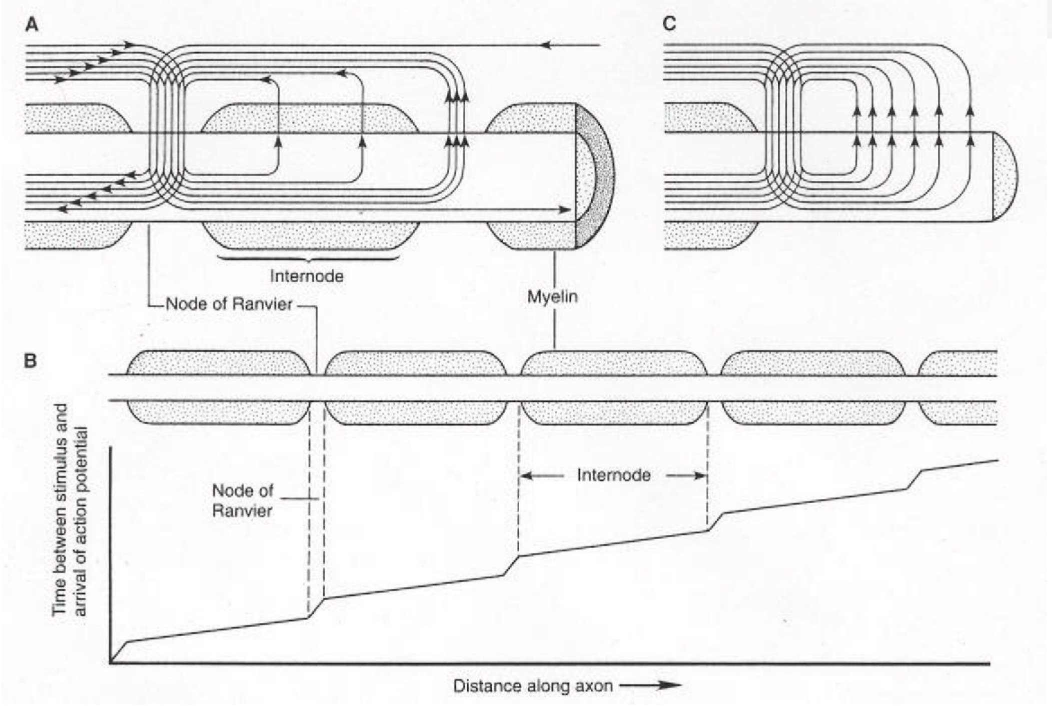


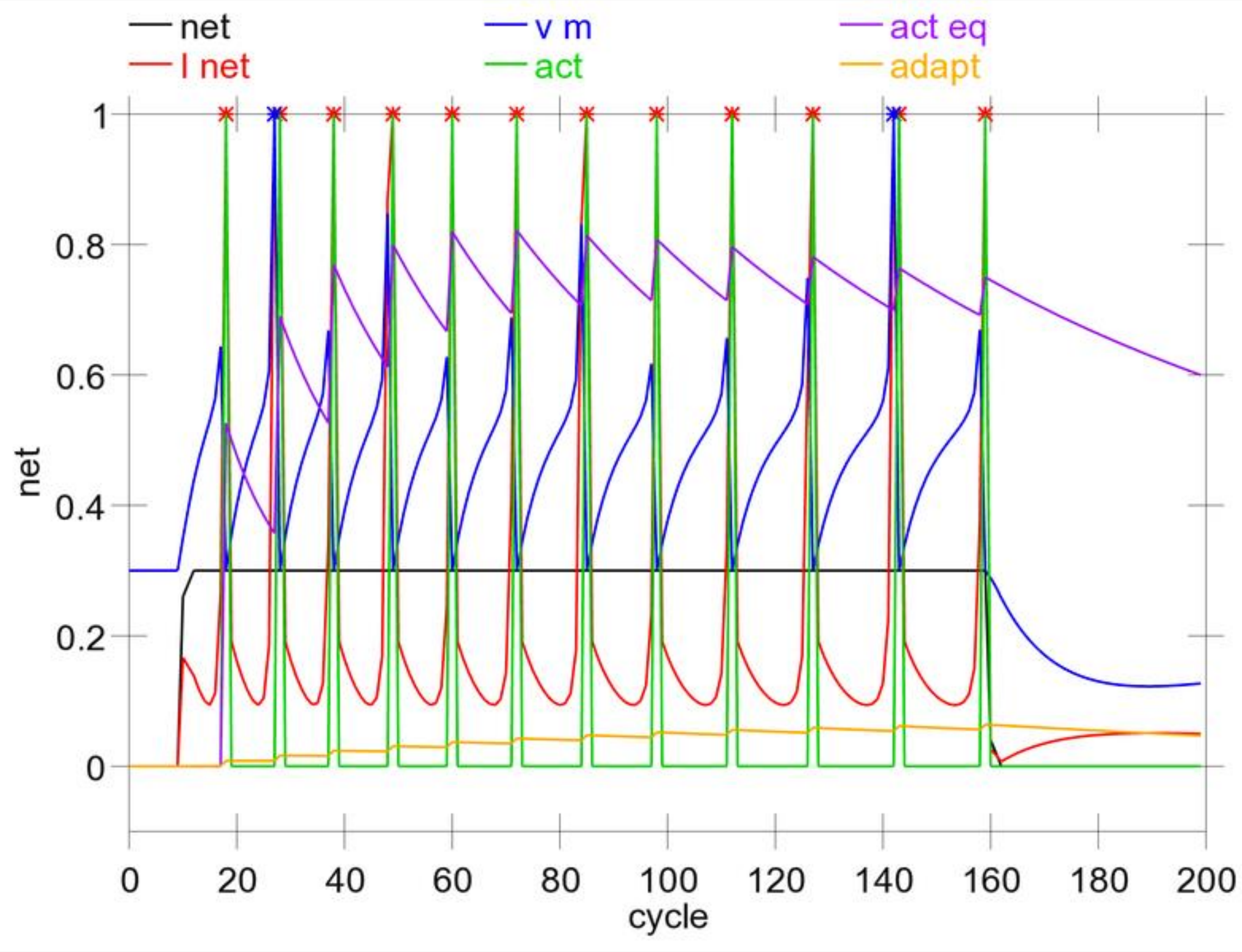
Cable Models

- String together the models to form a cable (signal drops exponentially)

- Voltage-gated sodium channels provide the active process needed to propagate signals.
- This is made efficient through *myelination* which insulates the cell between the *Nodes of Ranvier* (where the channels reside).



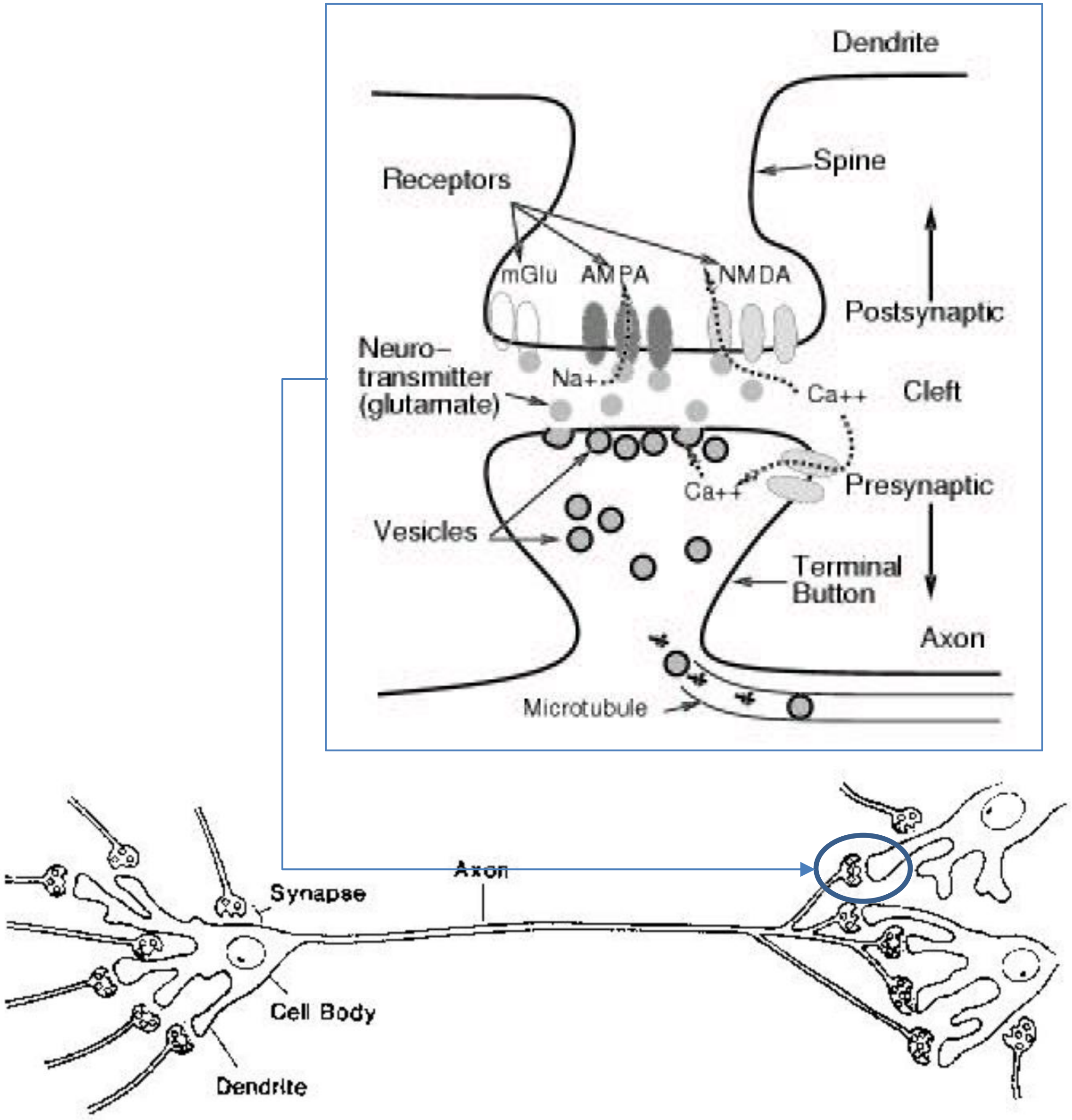




“Spiking” Neural Network Models

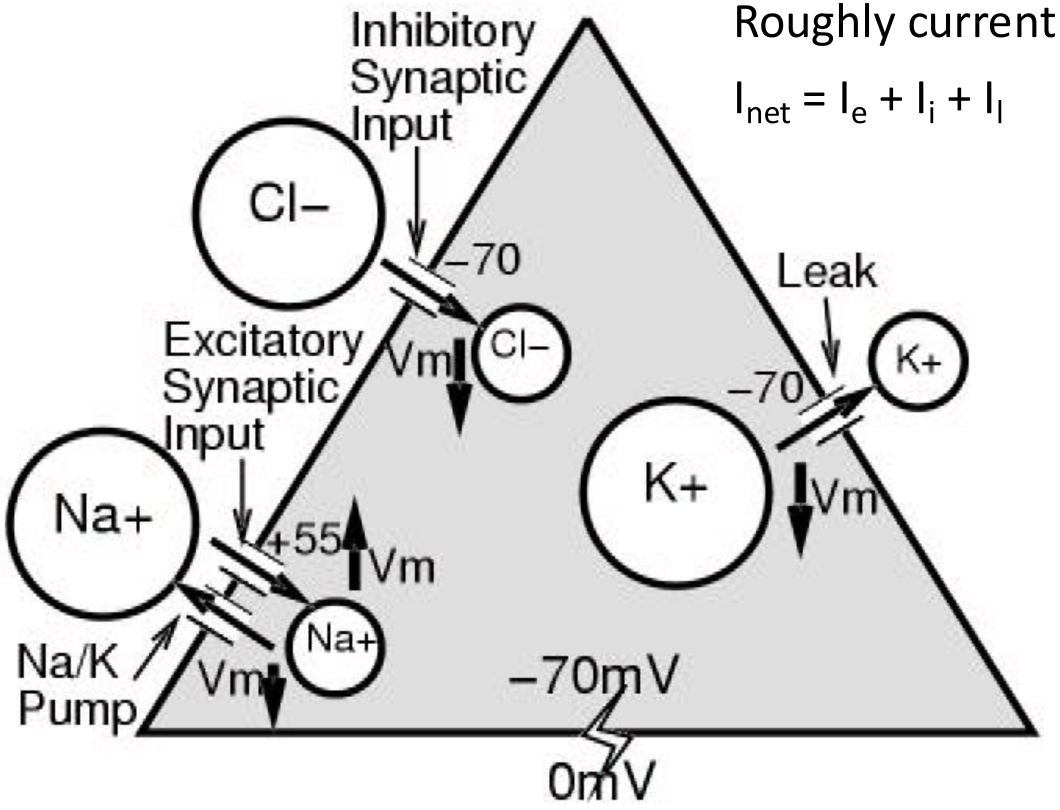
- Many modern artificial neural networks depend on precise timing and delivery of individual spikes to match performance compared to humans and animals: we will instead utilize the typical abstract model of this process throughout the course. However, we may see spiking models again later when we return to cognitive science…

Network Connection – The Synapse



- vesicles of neurotransmitter
- receptor sites: ionotropic & metabotropic
  - metabotropic glutamate (mGlu)
  - α-amino-3 hydroxy-5 methyl-4 isoxazole proprionic acid (AMPA)
  - N-methyl-D-aspartate (NMDA)
  - gamma-aminobutyric acid (GABA-A & GABA-B)

Dendritic Receptor Channels



Roughly current (I) for the cell:

$I_{net} = I_e + I_i + I_l$

Remember:

R = V / I

G = 1 / R

So…

I = GV

and

V=G/I

What does a neuron do?

- Neurons can basically be viewed as *dedicated feature detectors*.
- While this viewpoint is appropriate for all cells involved in all aspects of cognition, it is often difficult to characterize *what feature* a given neuron is detecting.

P(h|d) =

P(d|h)P(h) / P(d) =

P(d|h)P(h) /
[P(d|h)P(h) + P(d|h')P(h')]

$$
\begin{aligned}
V_m &= \frac{g_e(t)\bar{g}_e E_e + g_i(t)\bar{g}_i E_i + g_l(t)\bar{g}_l E_l}{g_e(t)\bar{g}_e + g_i(t)\bar{g}_i + g_l(t)\bar{g}_l} \\
&= \frac{g_e(t)\bar{g}_e}{g_e(t)\bar{g}_e + (\ g_i(t)\bar{g}_i + g_l(t)\bar{g}_l\ )} \\
&\approx \frac{P(d|h)\ P(h)}{P(d|h)\ P(h)\ +\ P(d|\bar{h})\ P(\bar{h})}
\end{aligned}
$$

Additional Information

- You can find additional information on neurons and their relationships to their artificial cousins in:
  - Computational Cognitive Neuroscience (O'Reilly et al., 2012)
- Free online: https://compcogneuro.org/
- Chapter on Neurons: https://github.com/CompCogNeuro/book/blob/main/chapter-02.md