Artificial Neural Networks

Artificial Neural Networks

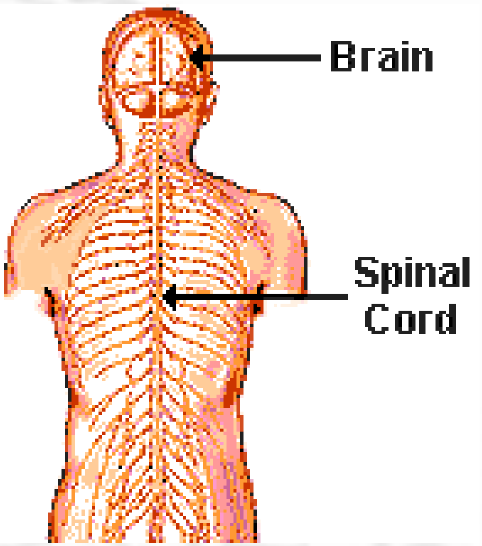
- Connectionist, PDP, etc. models
- A biologically-inspired approach for
  - intelligent computing machines
  - massive parallelism
  - distributed computing
  - learning, generalization, adaptivity
  - Tolerant of fault, uncertainty, imprecise info

Compared to Von Neumann

	Von Neumann computer	Biological neural systems
Processor	complex, high speed, few	simple, low speed, many
Memory	separate from processor, non-content addressable	integrated into processor, content addressable
Computing	centralized, sequential stored programs	distributed, parallel self-learning
Reliability	vulnerable	fault tolerant

Anatomy and Pathway

- Brain
  - Cerebrum
    - Frontal, parietal, temporal and occipital lobes
  - Cerebellum
  - Brainstem
- Spinal Cord
  - Housed in vertebral column
- Receptors to afferent neurons to brain (neocortex) to efferent neurons to effectors



Central Nervous System

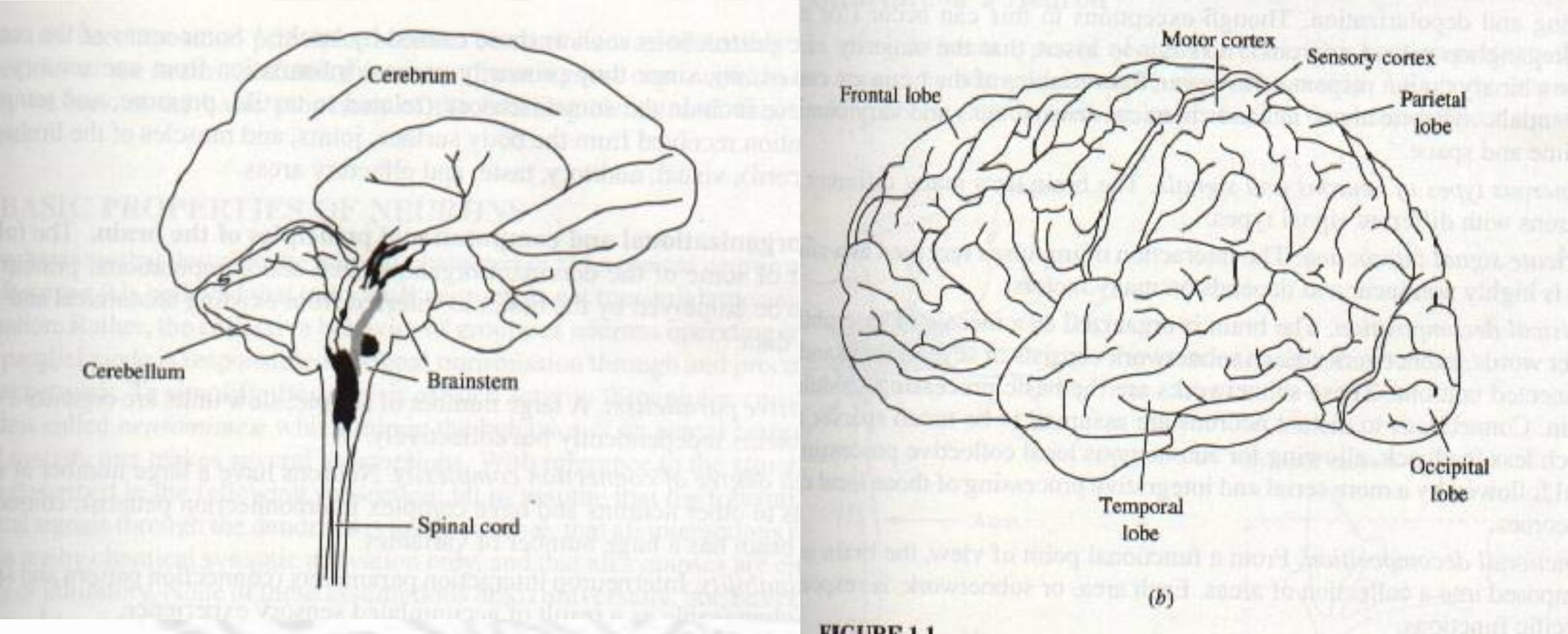


FIGURE 1.1
Two views of the human brain: (*a*) cross-sectional side view showing the major parts: cerebrum, cerebellum, and brainstem; (*b*) view showing cortex and lobes.

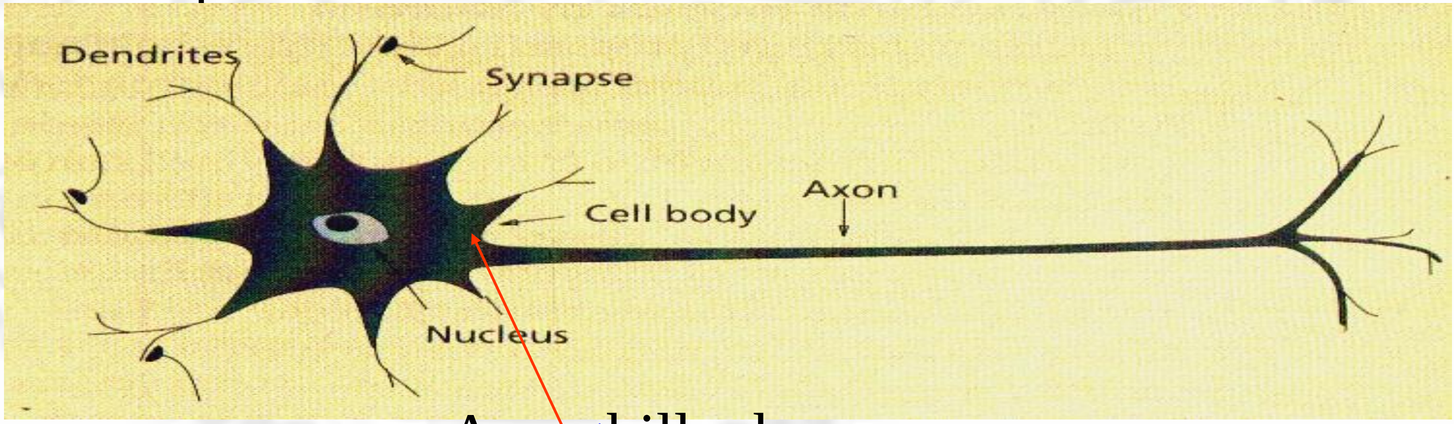
CNS

- Neocortex
  - Gray matter (surface layer of cerebrum)
  - White matter (connection fiber)
  - A crumbled paper analogy (folded and refolded many times to fit)
  - Regions (lobes) for different functions
- It is the
  - Activities of neurons (resting or depolarization)
  - Topology of the connection
  - Strength and direction (forward & backward) of connection

In the cortex that defines intelligence

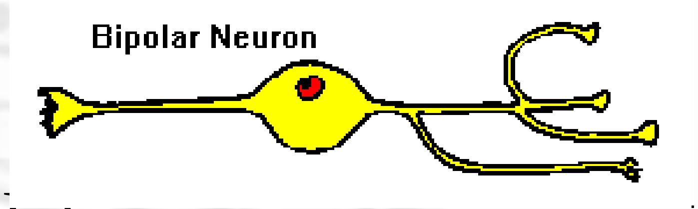
Biological Neural Networks

- soma (cell body)
- dendrites (receivers)
- axon (transmitters)
- synapses (connection points, axon-soma, axon-dendrite, axon-axon)
- Chemicals (neurotransmitters)
- 10^{11} neurons
- each makes about $10^3 \sim 10^4$ connections
- with an operating speed of a few milliseconds
- one-hundred-step rule



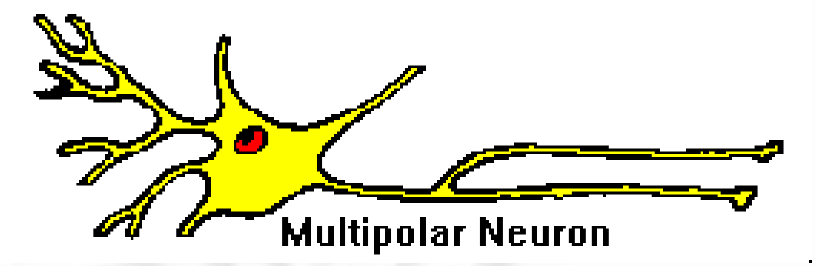
Axon hillock

That is not Quite Right



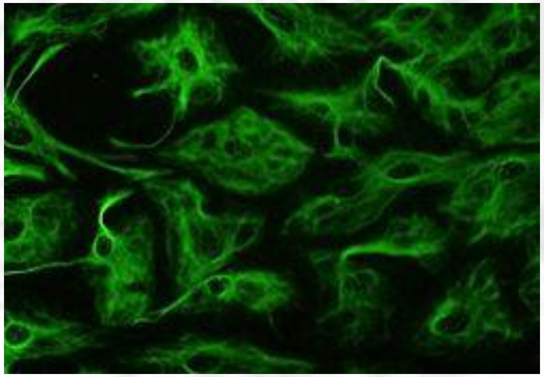
- Bipolar
  - E.g., found in eyes (retinal cells)
- Pseudounipolar
  - Two axons one to spinal cord one to skin and muscle

- Multipolar
  - Axons + dendrites



Glia

- Support cells
  - Clean up brain debris
  - Transport nutrients to neurons
  - Hold neurons in place
  - Digest dead neurons
  - Regulate content of extracellular space
- Insulation (myelin)
- Difference
  - Star shaped – no axons
  - No action potential
  - No synapses
  - A lot more (10 to 50 times more)



Signal Generation

- Resting potential
  - Charge difference across neuron membrane approximately –70mV
- Graded potential
  - Stimulus across synapses of post-synaptic neuron
- Action potential
  - If accumulation of graded potential across neuron membrane over a short period of time is higher than ~15mV, action potential is generated and propagated across axon
  - Same form and amplitude regardless of stimulus, signal by frequency rather than amplitude

Signal Generation

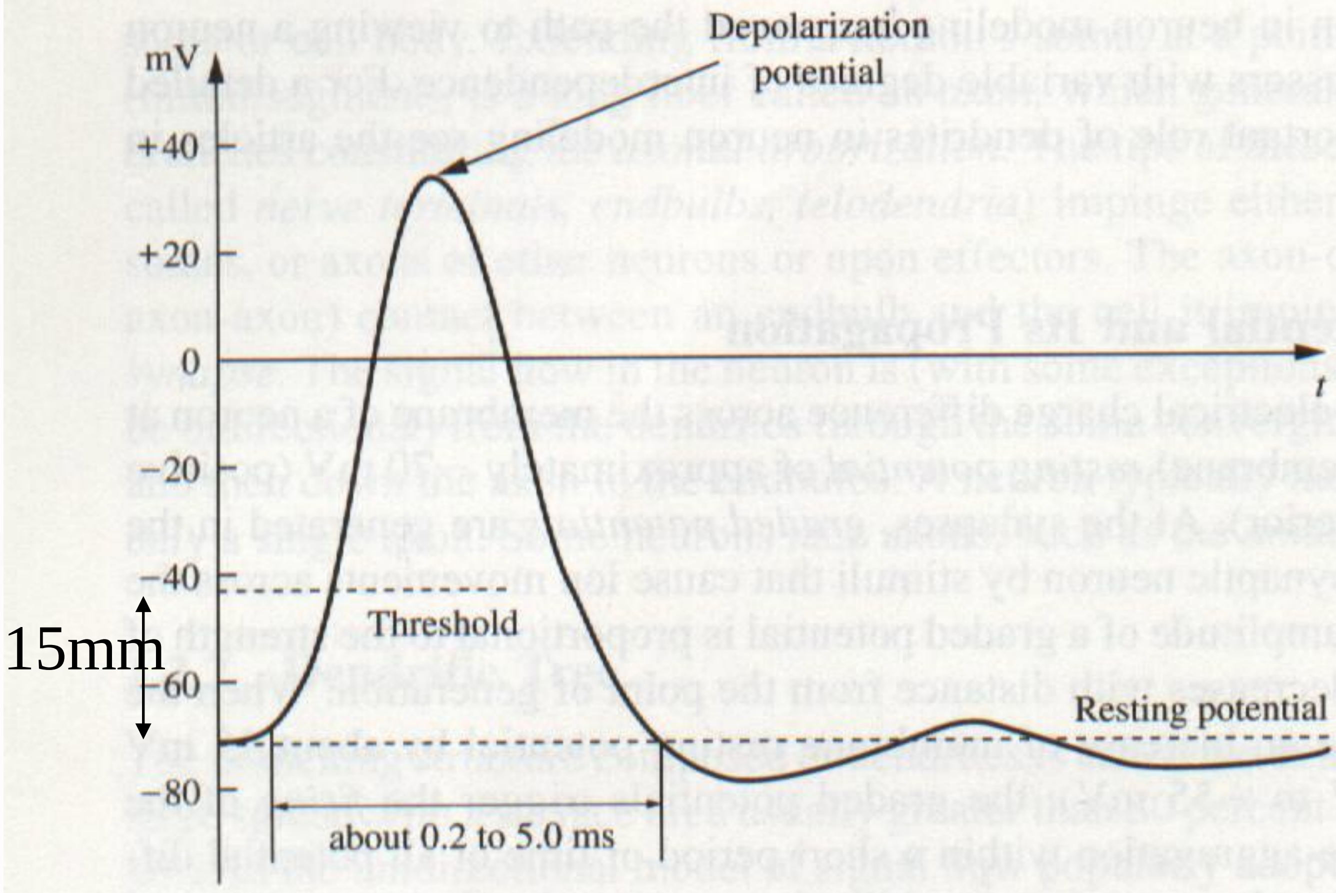
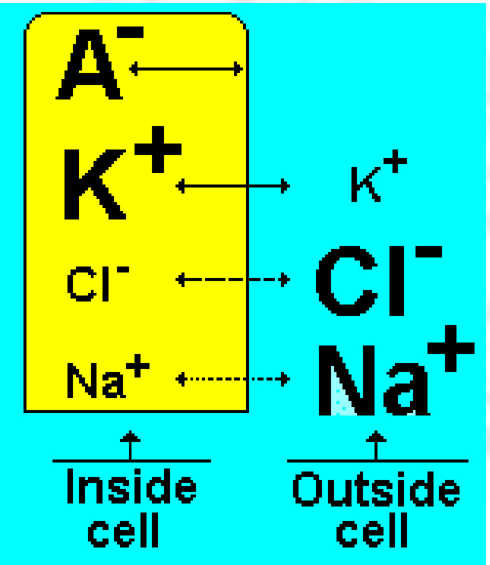


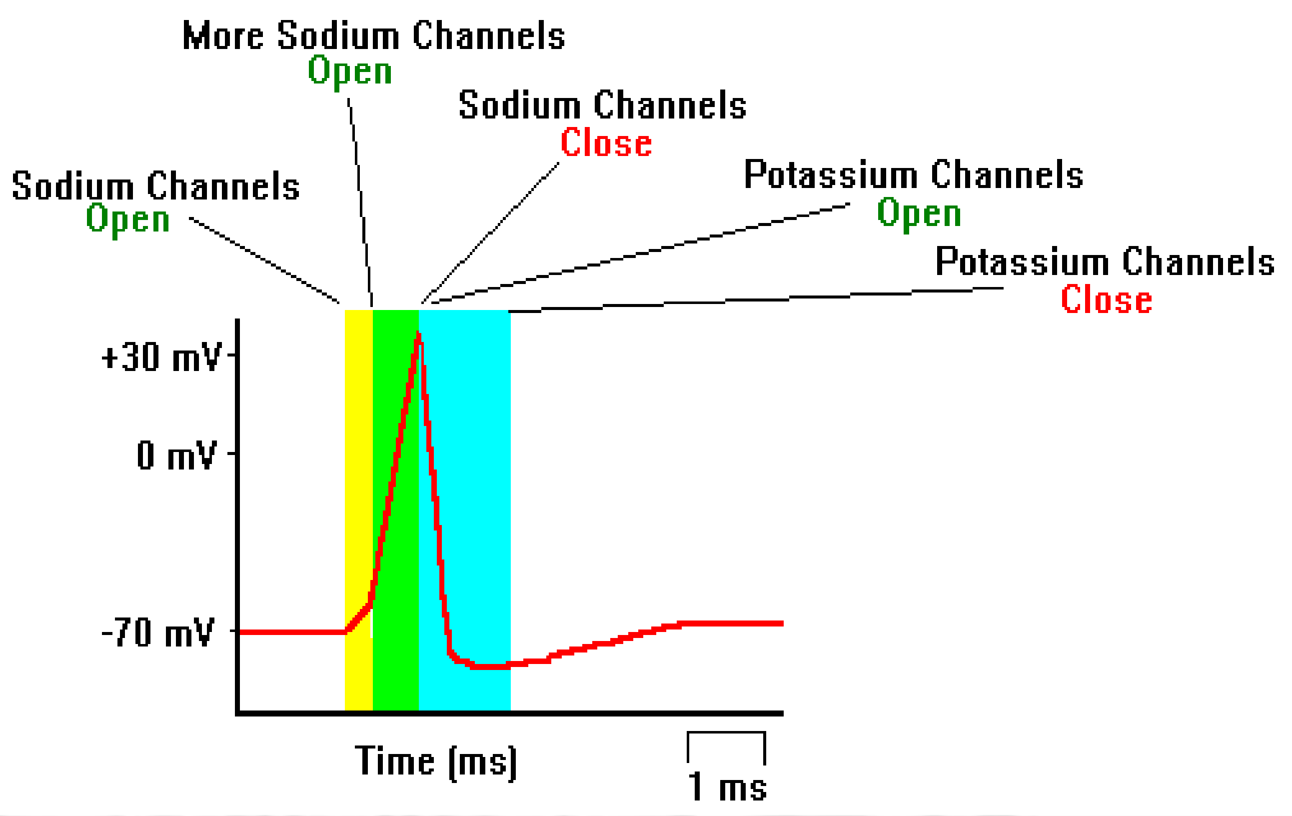
FIGURE 1.3
An action potential.

Signal Generation

- Resting
  - A-: protein
  - K+: potassium
    - Can go out
  - Cl-: cloride
  - Na+: sodium
    - Cannot go out



- Excited



- Think about electrical circuit
- Flow one way due to charge (concentration) difference
- Flow the other way using cell's ion pumps (battery)

Signal Flow in Dendritic Trees and Axons

- Flow is usually one dimensional
- Longitudinal flow (little transverse flow) with no loss (active transmission line)
- Myelindated
  - Wrapped in sheath of myelin, 100m/s
- Unmyelinated
  - 1m/s

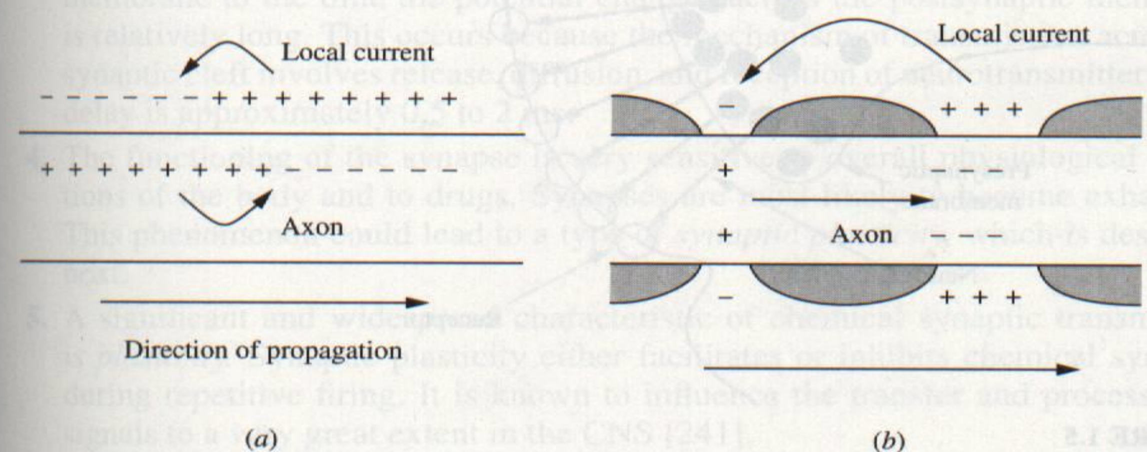
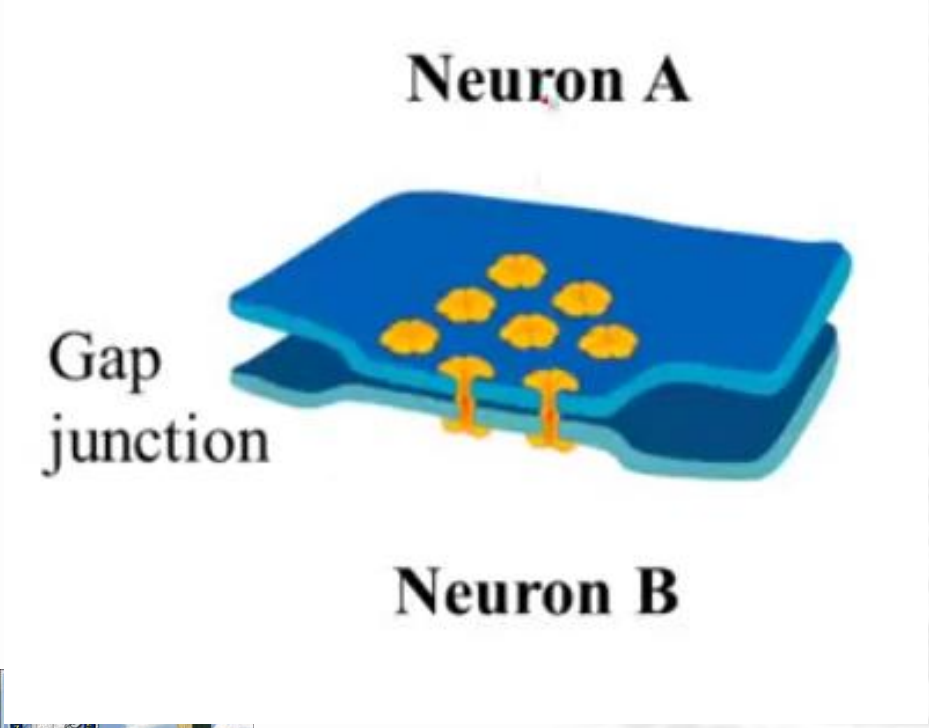
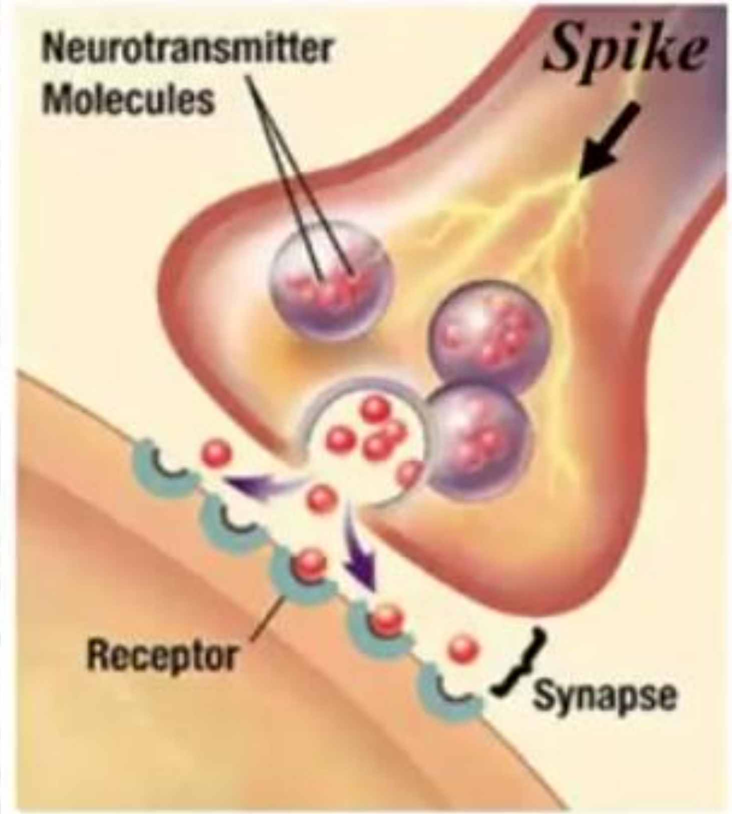


FIGURE 1.4
Local currents in an axon. During the generation of an action potential, the cell membrane becomes very permeable to Na^+ (positively charged sodium) ions, which flow inside, causing the membrane potential to rise. After a short while, the membrane becomes impermeable to Na^+ ions but permeable to K^+ (positively charged potassium) ions, which move outside to restore the membrane to its resting potential. (*a*) Unmyelinated axon; (*b*) myelinated axon.

synapses

- Electrical
  - Fixed, simultaneous control
- Chemical
  - Pattern and strength can be learned and adjusted





Signal Flow Across Synapses

- Action potential causes release of neurotransmitters from synaptic vesicles
- ~100 different neurotransmitters, e.g., dopamine, serotonin, and acetylcholine
- The release, diffusion and reception of neurotransmitters cause delay of 0.5 to 2ms
- Synaptic plasticity: either facilitate or inhibit chemical synapses

Signal Flow Across Synapses

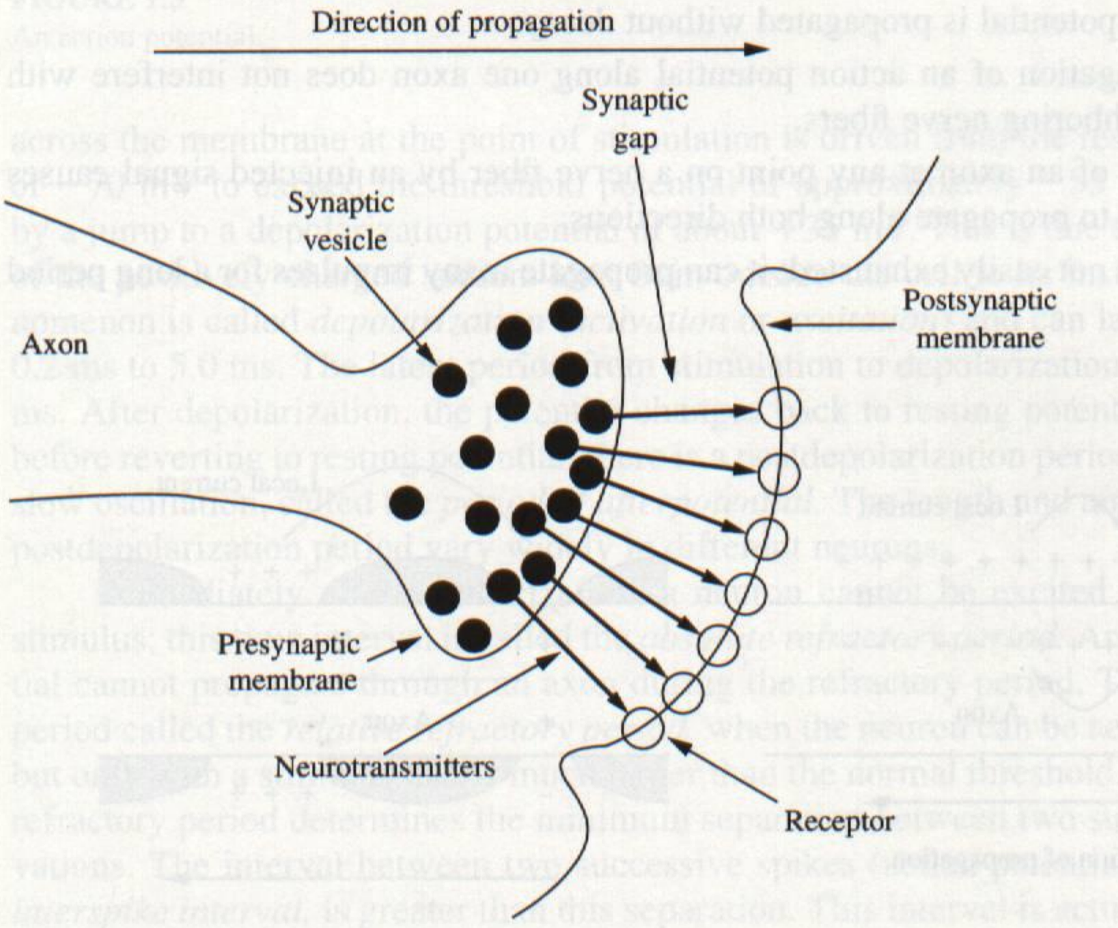


FIGURE 1.5
Transmission of presynaptic axon signal across synaptic gap (cleft) to postsynaptic neuron.

Connection Patterns

- Divergent (fan-out):
  - parallel processing, afferent neurons
- Convergent (fan-in):
  - Efferent neurons to effectors
- Chain and loop

Connection Patterns

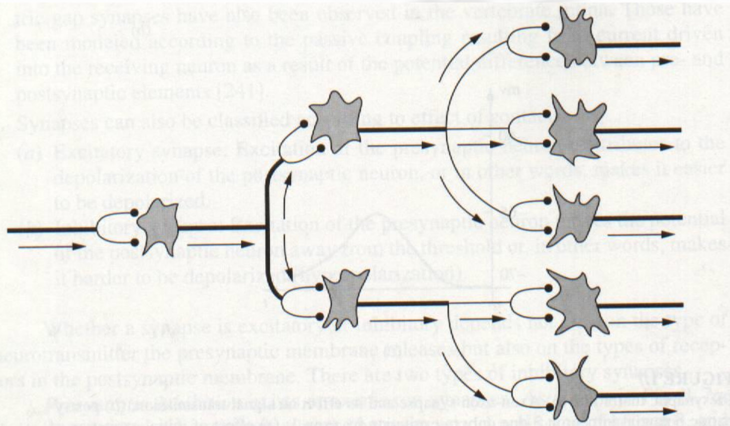


FIGURE 1.8
Divergent connections.

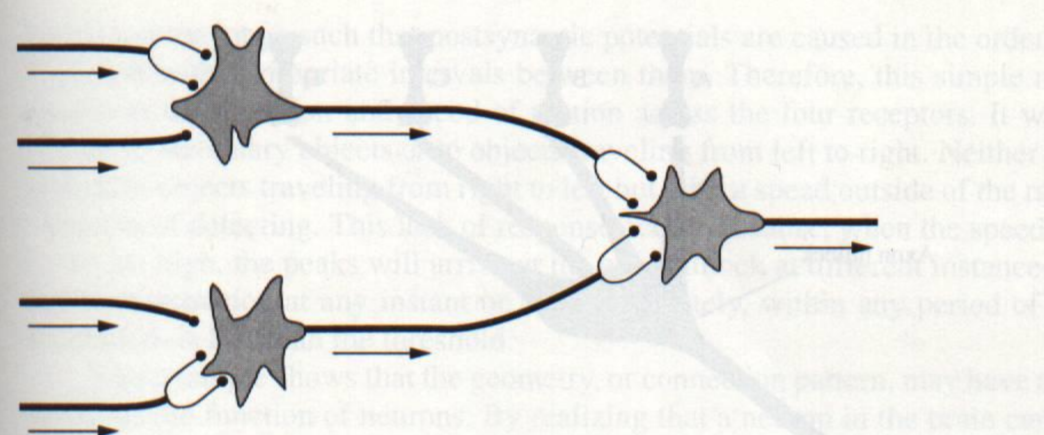


FIGURE 1.9
Convergent connections.

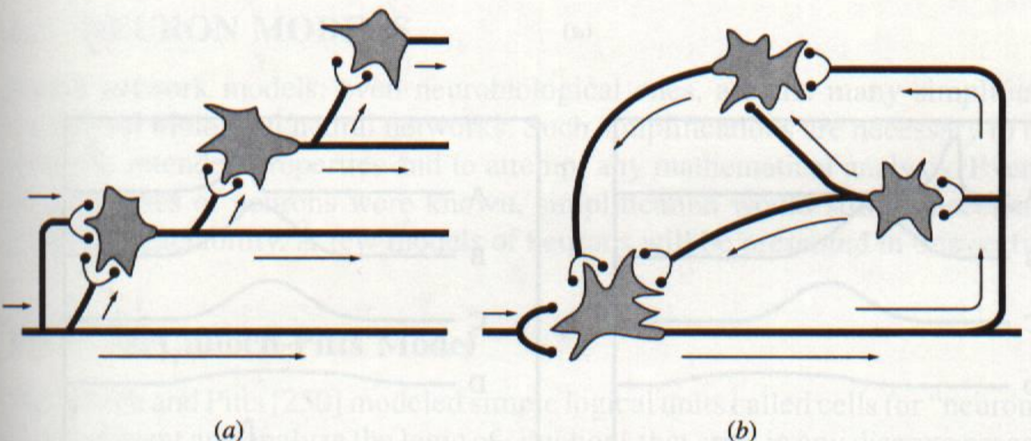
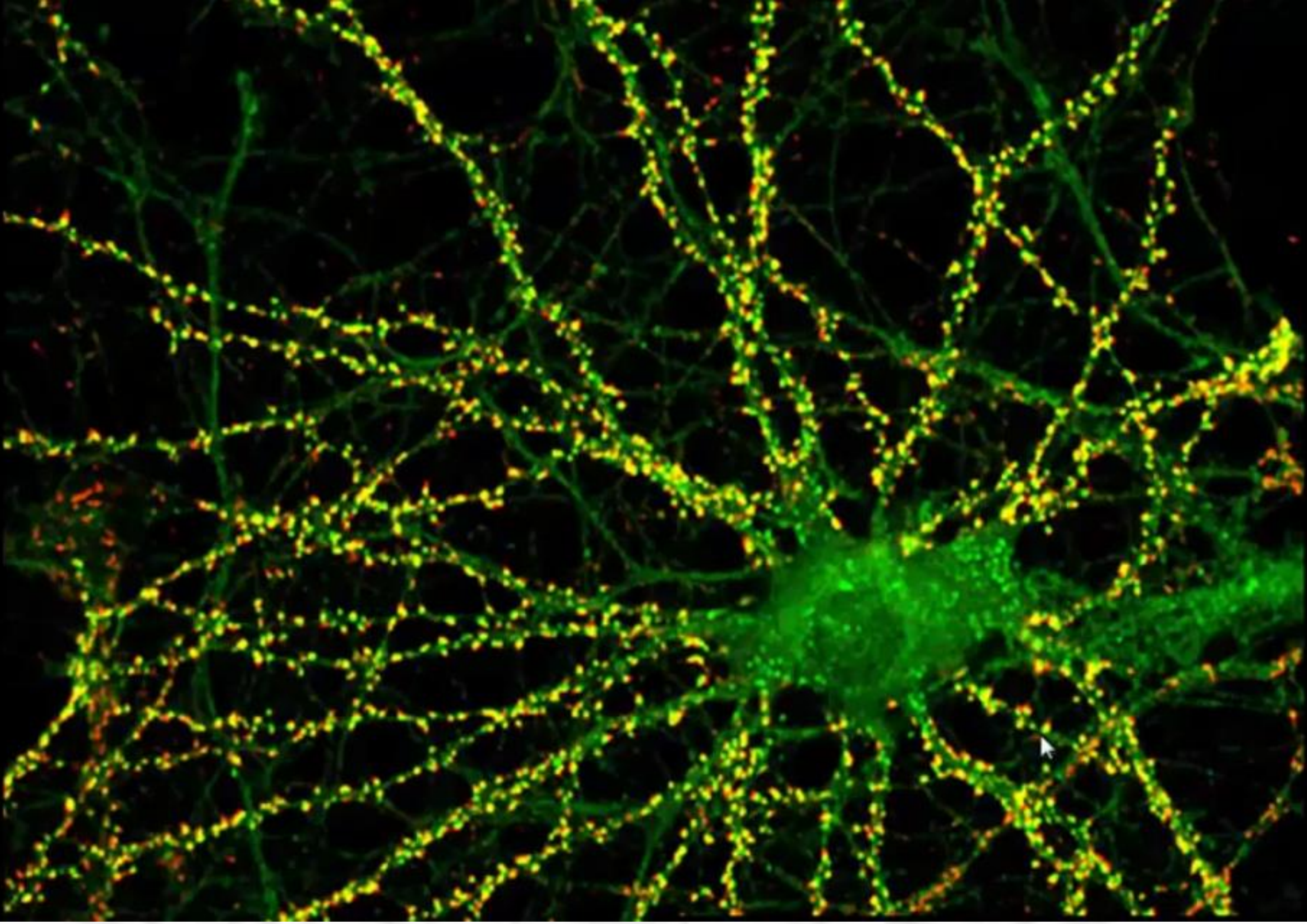
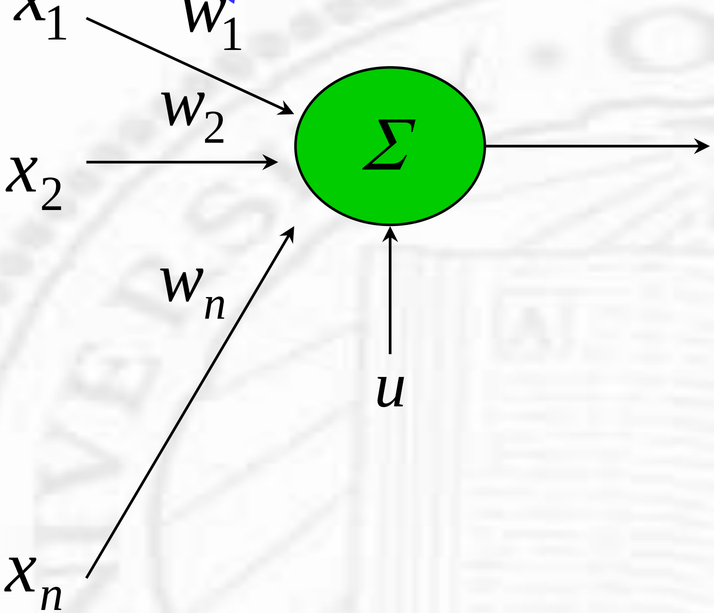


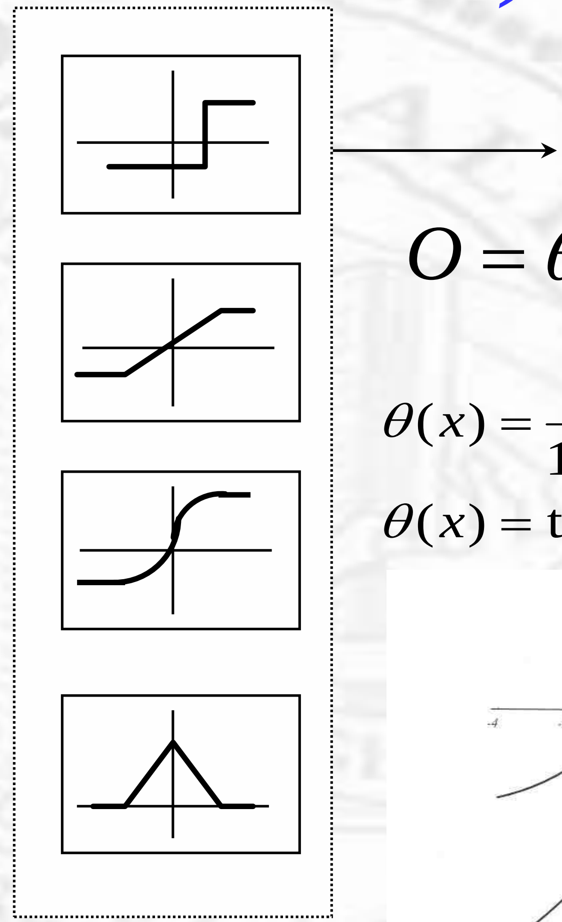
FIGURE 1.10
(*a*) Chain and (*b*) loop connections.



A single neurons with thousands of synapses (light yellow dots)

Computational Neuron Model (McCulloch and Pitts)

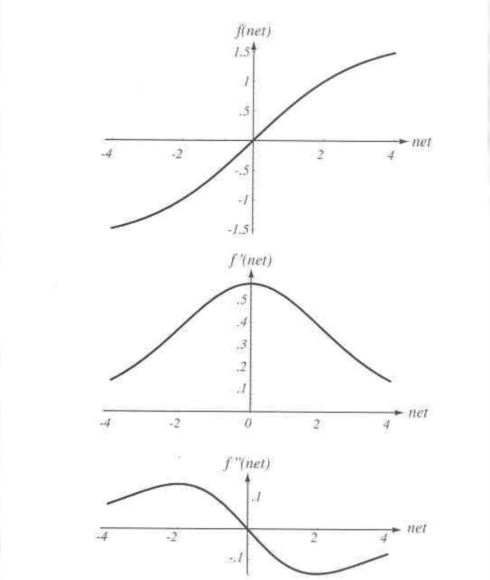




$$O = \theta\left(\sum_{i=1}^{n} w_i x_i - u\right)$$

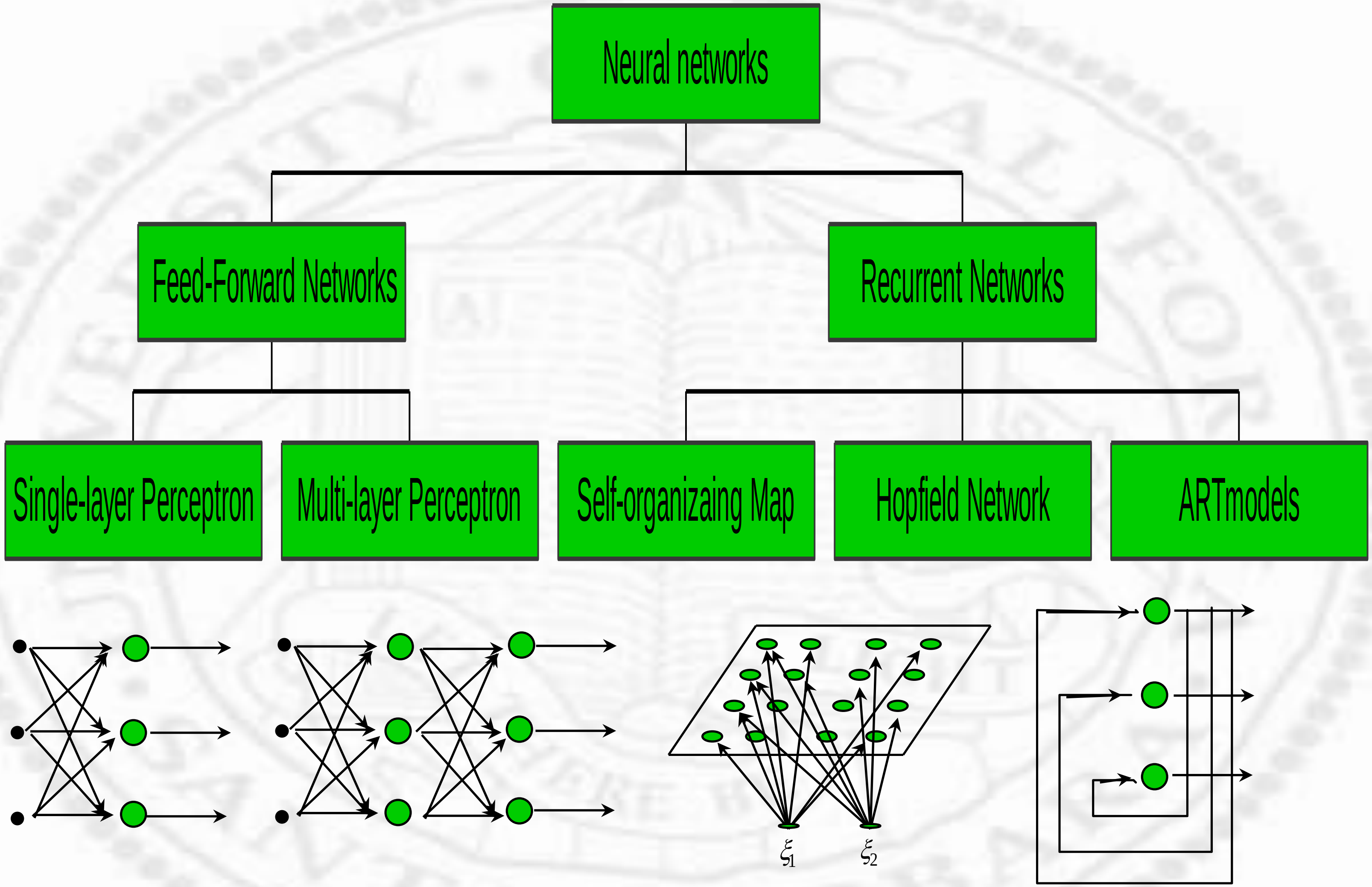
$$\theta(x) = \frac{1}{1+e^{-x}}, 0 \le \theta \le 1$$

$$\theta(x) = \tanh(x), -1 \le \theta \le 1$$



- time dependency?
- frequency response?

Computational Network Architecture



ANN Formulation

- Learning *rules* – Basic “workhorse” mechanism in adjusting weights of neurons
  - *Error-correcting learning (gradient descent)*
  - *Memory-based learning (nearest neighbor)*
  - *Hebbian learning (mutual excitation)*
  - *Competitive learning (winner-take-all)*
  - *Boltzmann learning (statistical mechanics)*

ANN Formulation (cont.)

- Learning *paradigms* – the big picture
  - Supervised (learning w. a teacher): correct I/O association is provided
  - Unsupervised (learning w/o. a teacher): discover similarity, inherent structure, and interesting patterns
  - Delayed (learning w. a critic):
- Theory
  - capacity: how many patterns can be stored
  - sample complexity: how many training patterns
  - computational complexity: training time

Relation to Pattern Recognition

- Supervised mode
  - Single & multi-layer perceptrons for learning complicated decision boundaries
- Unsupervised mode
  - Competitive and self-organization maps for constructing clusters