Introduction to Artificial Neural Networks

Lecture 1:

Biological Neurons and Neural Networks

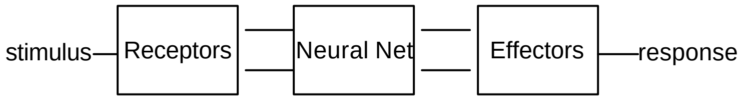
By: Ali Motie Nasrabadi

Outline

- The Nervous System
- Brains versus Computers: Some Numbers
- Structure of a Human Brain
- Basic Components of Biological Neurons
- How a biological neuron works?
- Organization of Levels in Brains
- Brain plasticity
- History of Artificial Neural Network
- AI and NN

Human Brain

- **Human nervous system**

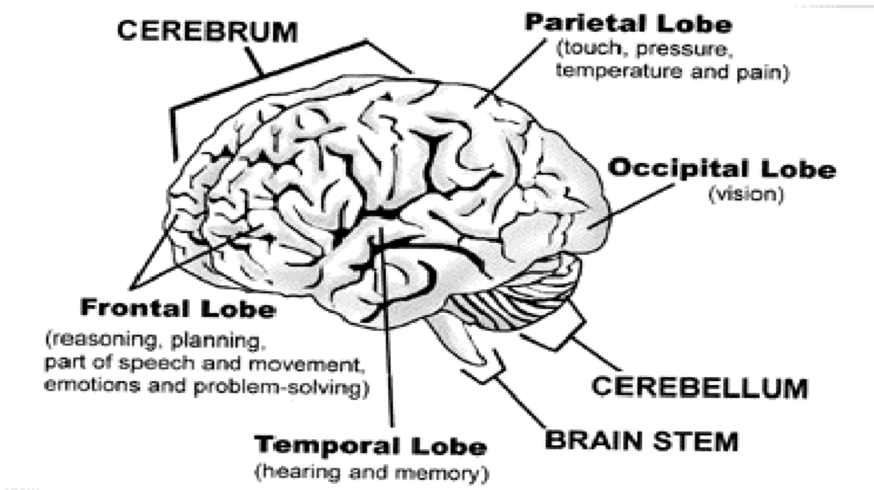


  - **10^{11} neurons in human cortex**
  - **60×10^{12} synaptic connections**
  - **10^{4} synapses per neuron**
  - **10^{-3} sec cycle time (computer : 10^{-9} sec)**
  - **energetic efficiency : 10^{-16} joules operation per second (computer : 10^{-6} joules)**

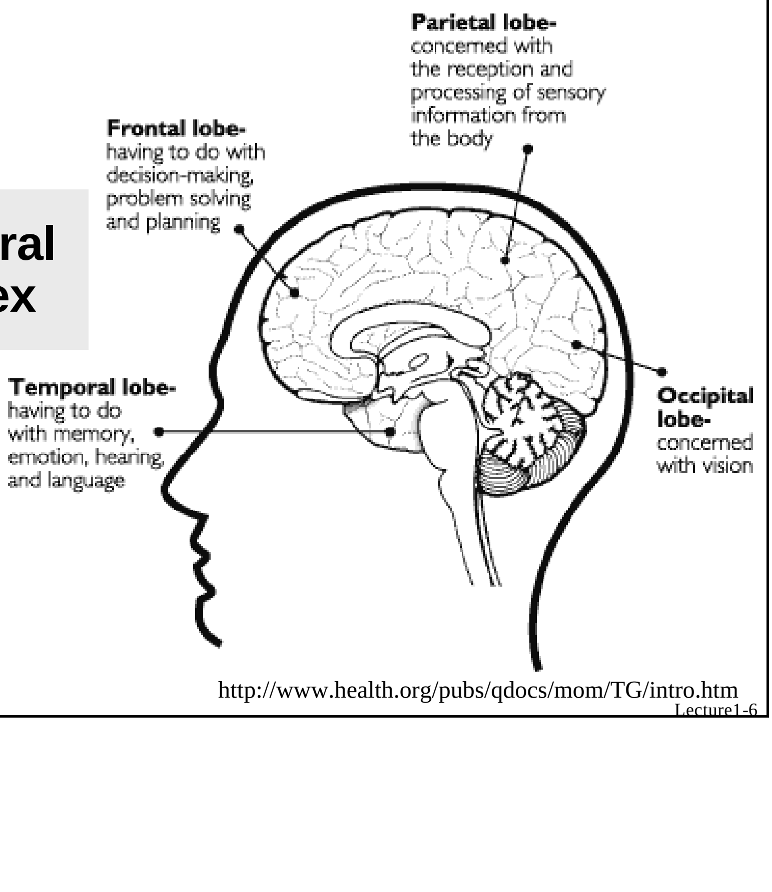
Brains versus Computers : Some numbers

- **There are approximately 10 billion neurons in the human cortex, compared with 10 of thousands of processors in the most powerful parallel computers.**
- **Each biological neuron is connected to several thousands of other neurons, similar to the connectivity in powerful parallel computers.**
- **Lack of processing units can be compensated by speed. The typical operating speeds of biological neurons is measured in milliseconds (10^{-3} sec), while a silicon chip can operate in nanoseconds (10^{-9} sec).**
- **The human brain is extremely energy efficient, using approximately 10^{-16} per operation per second, whereas the best computers today use around 10^{-6} joules per operation per second.**
- **Brains have been evolving for tens of millions of years, computers have been evolving for tens of decades.**

Structure of a Human Brain



Cerebral Cortex



http://www.health.org/pubs/qdocs/mom/TG/intro.htm

Slice Through a Real Brain

http://medlib.med.utah.edu/WebPath/HISTHTML/NEURANAT/NEURANCA.html

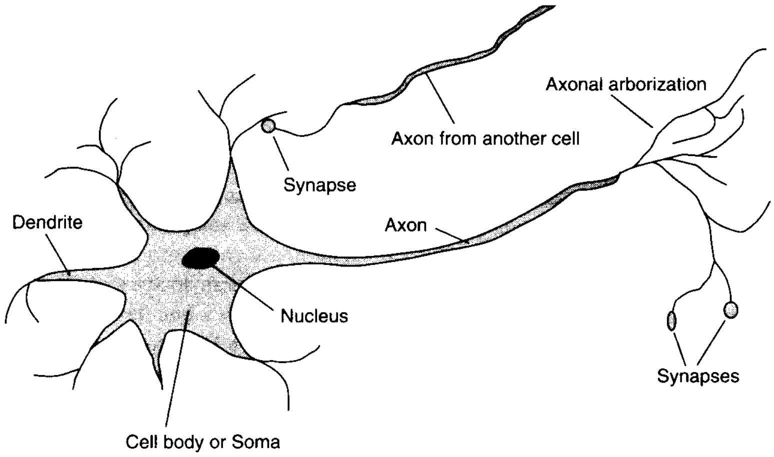
Lecture1-7

Human Brain

- Neuron structure
  - nucleus, cell body, axon, dendrite, synapses
- Neurotransmission
  - Neuron's output is encoded as a series of voltage pulses
    - called action potentials or spikes
  - by means of electrical impulse effected by chemical transmitter
  - Period of latent summation
    - generate impulse if total potential of membrane reaches a level : firing
  - excitatory or inhibitory
- Cerebral cortex
  - Areas of Brain for specific function

Lecture1-8

Basic Components of Biological Neurons



Excerpted from *Artificial Intelligence: Modern Approach*
by S. Russel and P. Norvig

Basic Components of Biological Neurons

- **The majority of** *neurons* **encode their activations or outputs as a series of brief electrical pulses (i.e. spikes or action potentials).**
- **The neuron's** *cell body (soma)* **processes the incoming activations and converts them into output activations.**
- **The neuron's** *nucleus* **contains the genetic material in the form of DNA. This exists in most types of cells, not just neurons.**
- *Dendrites* **are fibers which emanate from the cell body and provide the receptive zones that receive activation from other neurons.**
- *Axons* **are fibers acting as transmission lines that send activation to other neurons.**
- **The junctions that allow signal transmission between the axons and dendrites are called** *synapses*. **The process of transmission is by diffusion of chemicals called** *neurotransmitters* **across the synaptic cleft.**

Neurotransmission

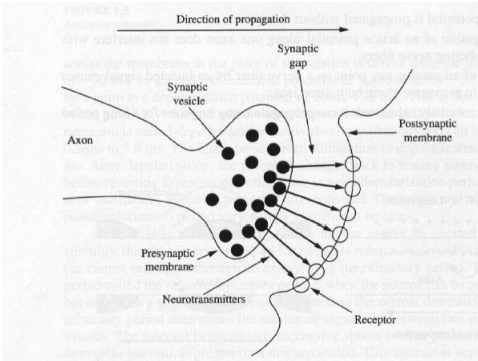
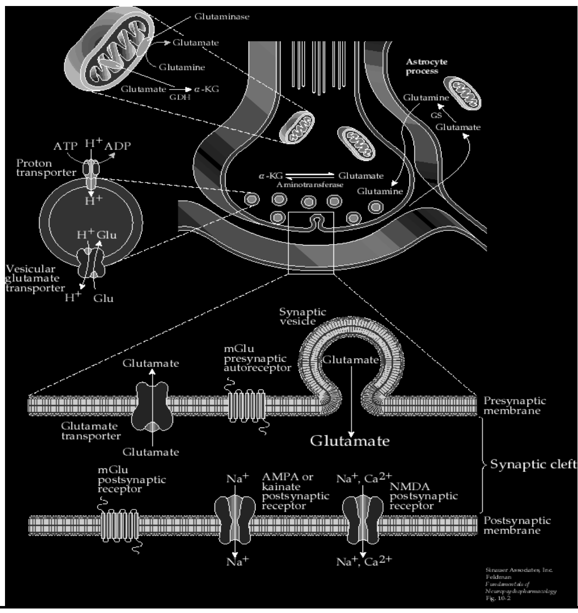


FIGURE 1.5
Transmission of presynaptic axon signal across synaptic gap (cleft) to postsynaptic neuron.

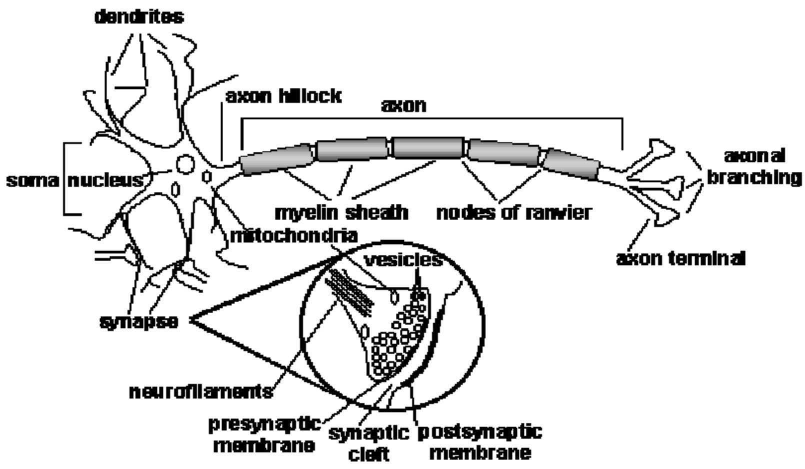
Lecture1-11

Neurotransmission



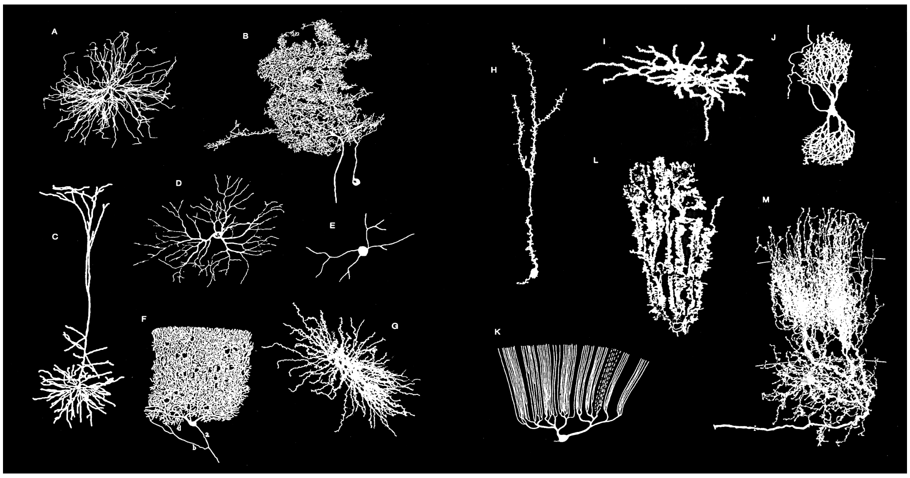
Lecture1-12

Schematic Diagram of a Biological Neuron



Lecture1-13

Anatomical Diversity



Lecture1-14

How a biological neuron works?

- **Signals are transmitted between neurons by electrical pulses (action-potentials or 'spike' trains) traveling along the axon.**
- **These pulses impinge on the afferent neuron at terminals called *synapses*.**
- **These are found principally on a set of branching processes emerging from the cell body (soma) known as *dendrites*.**
- **Each pulse occurring at a synapse initiates the release of a small amount of chemical substance or *neurotransmitter* which travels across the synaptic cleft and which is then received at post-synaptic receptor sites on the dendritic side of the synapse.**
- **The neurotransmitter becomes bound to molecular sites here which, in turn, initiates a change in the dendritic membrane potential.**

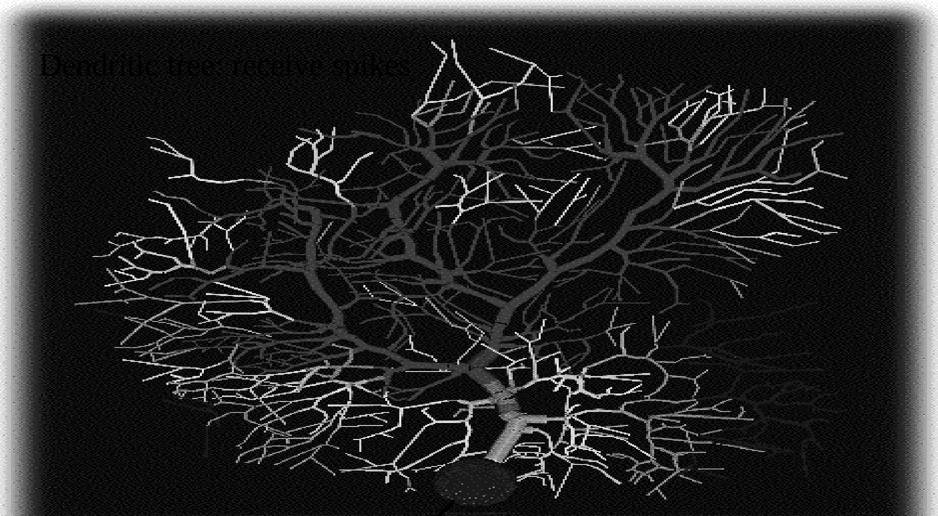
How a biological neuron works?

- **This post-synaptic-potential (PSP) change may serve to increase (hyperpolarize) or decrease (depolarize) the polarization of the post-synaptic membrane. In the former case, the PSP tends to inhibit generation of pulses in the afferent neuron, while in the latter, it tends to excite the generation of pulses.**
- **The size and type of PSP produced will depend on factors such as the geometry of the synapse and the type of neurotransmitter.**
- **Each PSP will travel along its dendrite and spread over the soma, eventually reaching the base of the axon (axon-hillock).**
- **The afferent neuron *sums* or *integrates* the effects of thousands of such PSPs over its dendritic tree and over time.**
- **If the integrated potential at the axon-hillock exceeds a threshold, the cell `fires' and generates an action potential or spike which starts to travel along its axon.**
- **This then initiates the whole sequence of events again in neurons contained in the efferent pathway.**

How a biological neuron works?

- **If the integrated potential at the axonhillock exceeds a threshold, the cell 'fires' and generates an action potential or spike which starts to travel along its axon.**
- **This then initiates the whole sequence of events again in neurons contained in the efferent pathway.**

Lecture1-17

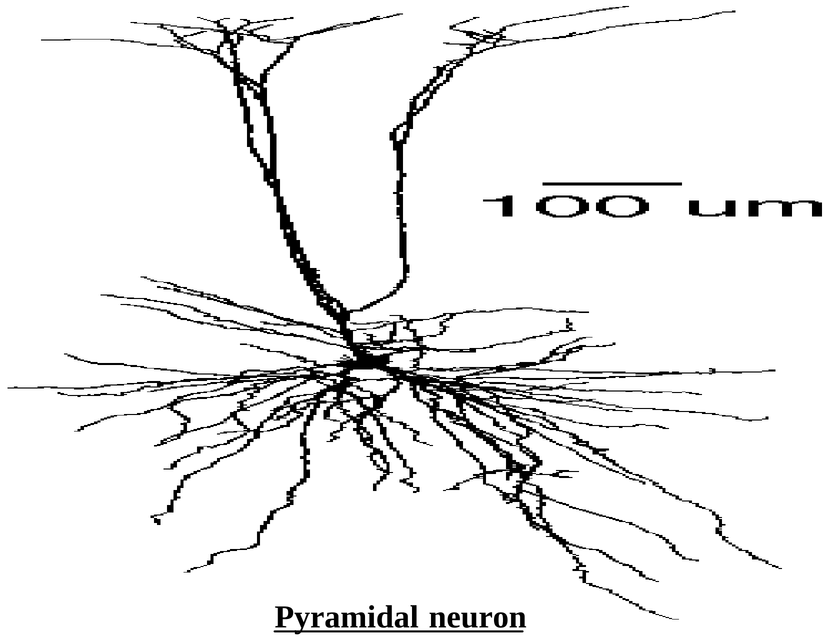


Soma: generate spikes

Axon: sent out spikes

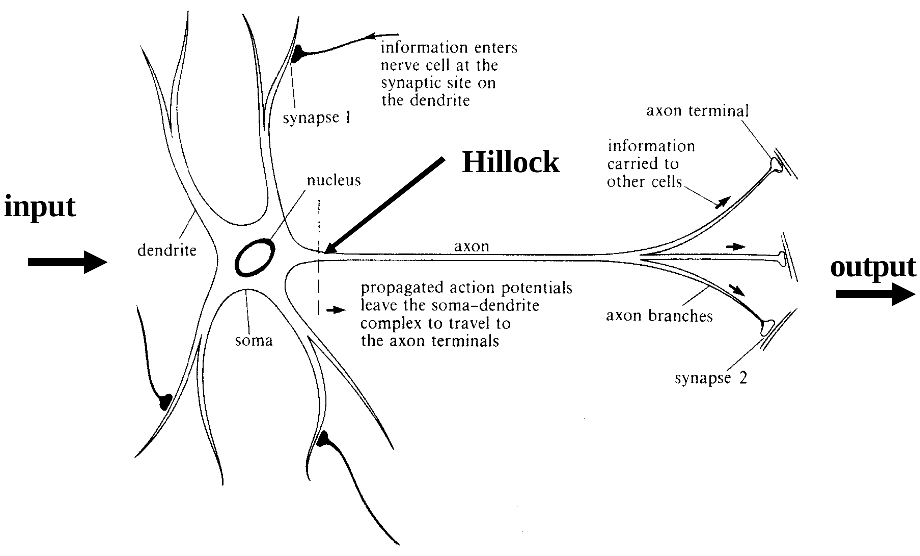
Purkinje neuron

Lecture1-18



Pyramidal neuron

Neurons



Lecture1-20

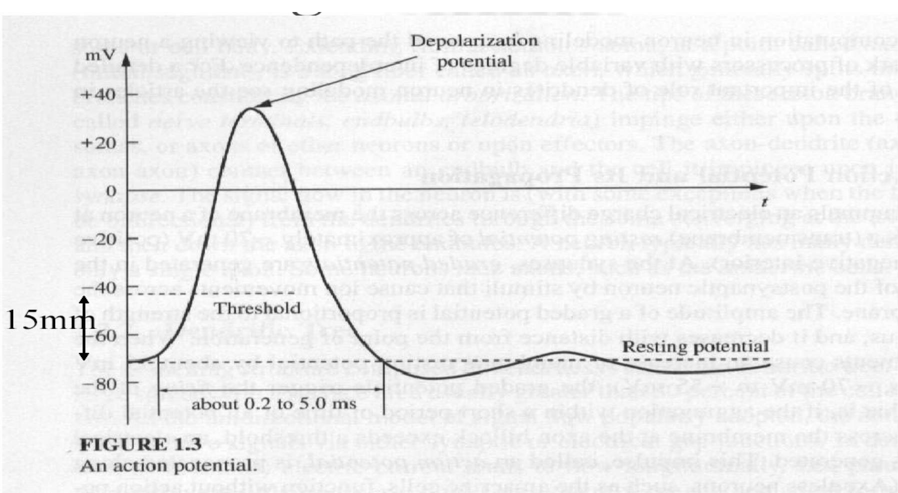
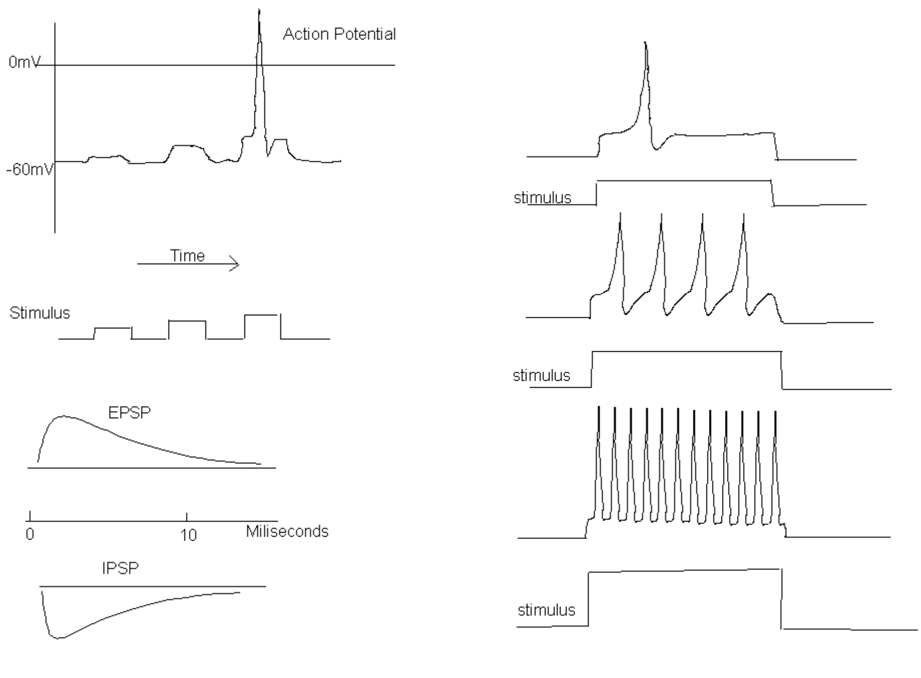


FIGURE 1.3
An action potential.

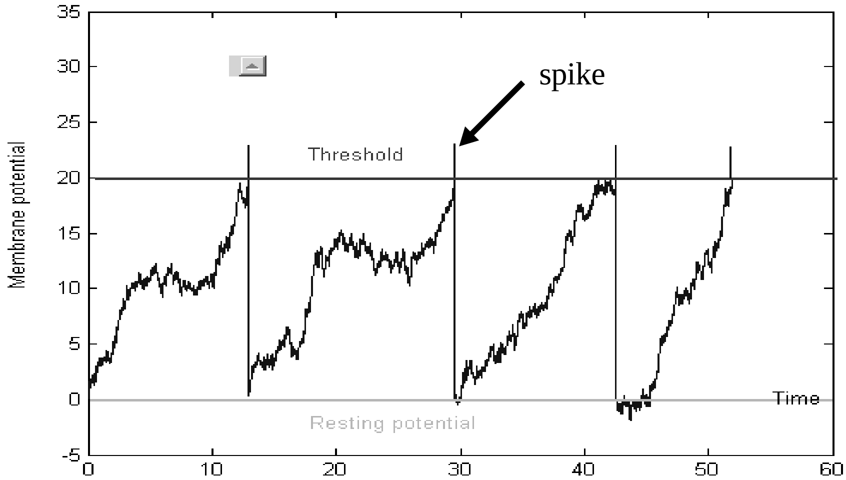
Lecture1-21



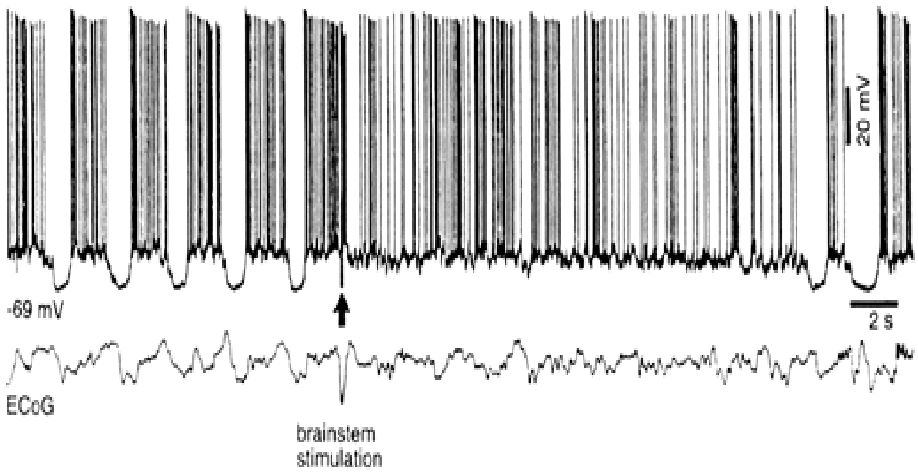
Lecture1-22

Single neuron activity

•If you measure the ***membrane potential*** of a neuron and print it out on the screen, it looks like (from time 0 to 60 minutes)

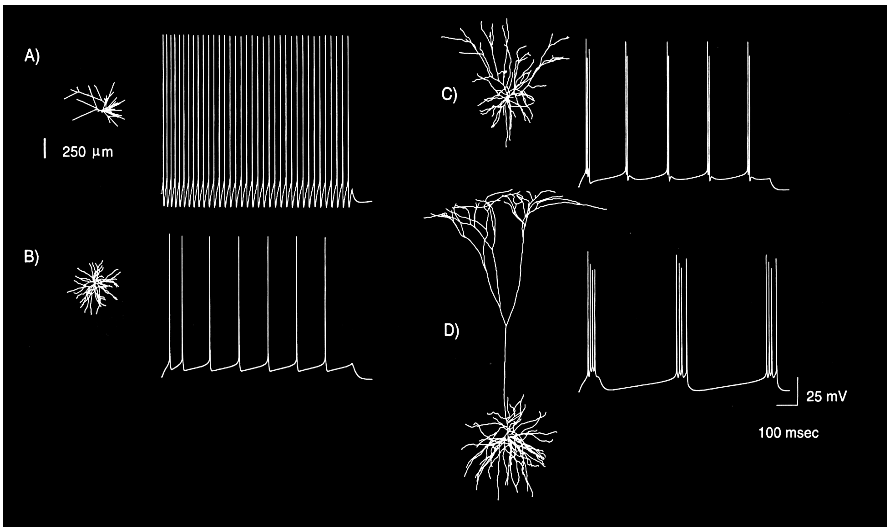


- **Only spikes are important since other neurons receive them (signals)**
- **Neurons communicate with spikes**
- **Information is coded by spikes**
- **If we can manage to measure the spiking time, we decipher how the brain works**



Recording from a real neuron: membrane potential

Divergent Single neuron activity



Lecture1-26

Connection Patterns

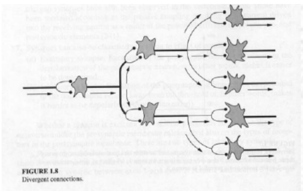


FIGURE 1.8
Divergent connections.

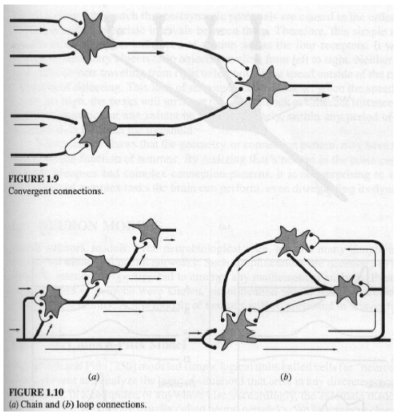


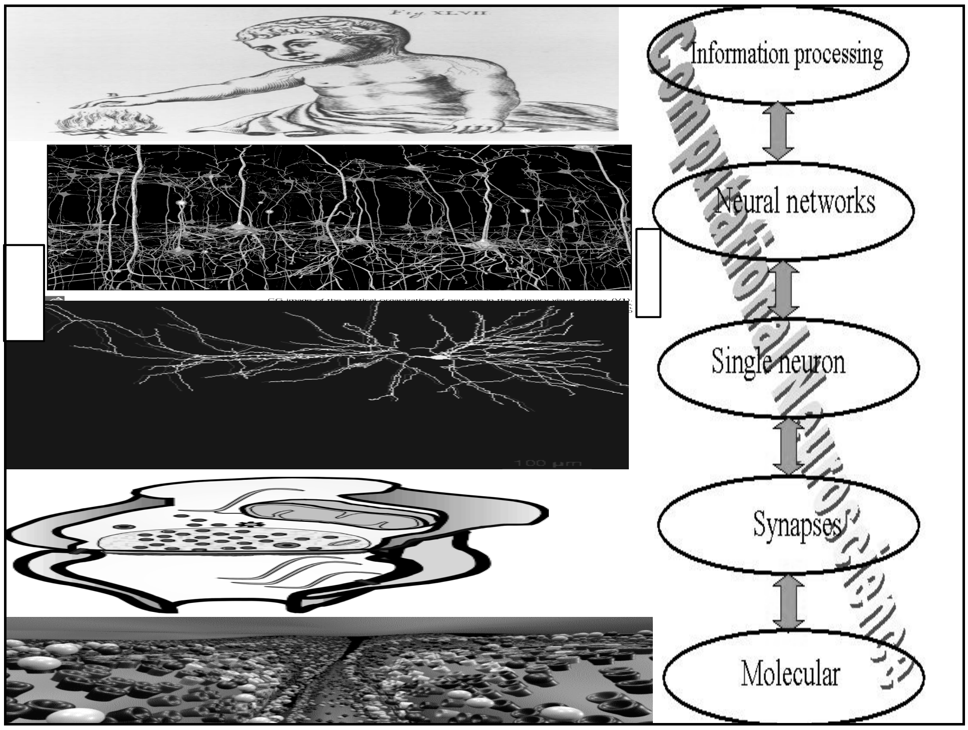
FIGURE 1.9
Convergent connections.

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

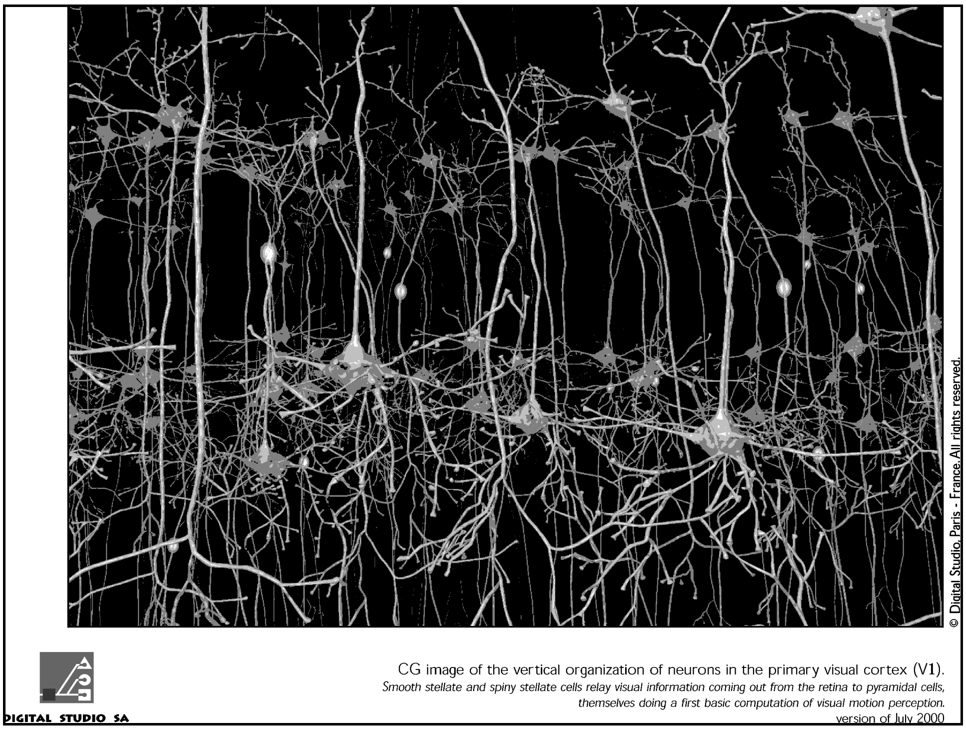
Organization of Levels in Brains

- Central Nervous sys
- Interregional circuits
- Local circuits
- Neurons
- Dendrite tree
- Neural microcircuits
- Synapses
- Molecules

- **map into cerebral cortex**
  pathways, columns, topographic maps; involve multiple regions
- **neurons of similar and different properties, 1 *mm* in size, localized region in the brain**
- **100μm in size, contains several dendrite trees**
- **Assembly of synapses (silicon chip), milisecond, μm**
- **most fundamental level (transistor)**



Information processing
Neural networks
Single neuron
Synapses
Molecular
Computational Neuroscience



© Digital Studio, Paris - France. All rights reserved.
CG image of the vertical organization of neurons in the primary visual cortex (V1).
Smooth stellate and spiny stellate cells relay visual information coming out from the retina to pyramidal cells, themselves doing a first basic computation of visual motion perception.
version of July 2000
DIGITAL STUDIO SA

Lecture1-31

Brain plasticity

- **At the early stage of the human brain development (the first two years from birth) about 1 million synapses (hard- wired connections) are formed per second.**
- **Synapses are then modified through the learning process (plasticity of a neuron).**
- **In an adult brain plasticity may be accounted for by the following mechanisms:**
  - **creation of new synaptic connections between neurons,**
  - **deletion of existing synaptic connections**
  - **modification of existing synapses.**

Lecture1-32

History

Minsky & Papert(1969) ------ Perceptrons

Rosenblatt(1960) ------ Perceptron

Minsky(1954) ------ Neural Networks (PhD Thesis)

Hebb(1949) --------The organization of behaviour

McCulloch & Pitts (1943) -----neural networks and artificial intelligence were born

History

spiking neural networks

Vapnik (1990) ---support vector machine

Broomhead & Lowe (1988) ----Radial basis functions (RBF)

Linsker (1988) ----- Informax principle

Rumelhart, Hinton & Williams (1986) -------- Back-propagation

Kohonen(1982) ------ Self-organizing maps
Hopfield(1982) ------ Hopfield Networks

Minsky & Papert(1969) ------ Perceptrons

Rosenblatt(1960) ------ Perceptron

Minsky(1954) ------ Neural Networks (PhD Thesis)

Hebb(1949) --------The organization of behaviour

McCulloch & Pitts (1943) -----neural networks and artificial intelligence were born

History

spiking neural networks (Feng)

Vapnik (1990) ---support vector machine

Broomhead & Lowe (1988) ----Radial basis functions (RBF)

Linsker (1988) ----- Informax principle

Rumelhart, Hinton & Williams (1986) -------- Back-propagation

Kohonen(1982) ------ Self-organizing maps

Hopfield(1982) ------ Hopfield Networks

Minsky & Papert(1969) ------ Perceptrons

Rosenblatt(1960) ------ Perceptron

Minsky(1954) ------ Neural Networks (PhD Thesis)

Hebb(1949) --------The organization of behaviour

McCulloch & Pitts (1943) -----neural networks and artificial intelligence were born

Lecture1-35

History

spiking neural networks (Feng, Cogs)

Vapnik (1990) ---support vector machine

Broomhead & Lowe (1988) ----Radial basis functions (RBF)

Linsker (1988) ----- Informax principle

Rumelhart, Hinton & Williams (1986) -------- Back-propagation

Kohonen(1982) ------ Self-organizing maps

Hopfield(1982) ------ Hopfield Networks

Minsky & Papert(1969) ------ Perceptrons

Rosenblatt(1960) ------ Perceptron

Minsky(1954) ------ Neural Networks (PhD Thesis)

Hebb(1949) --------The organization of behaviour

McCulloch & Pitts (1943) -----neural networks and artificial intelligence were born

Lecture1-36

History of NNs - 1

- By the 1940s, neurophysiologists knew that the brain consisted of billions of intricately interconnected neurons
- The neurons all seemed to be basically identical
- The idea emerged that the complex behaviour and power of the brain arose from the connection scheme
- This led to the birth of connectionist approach to the explanation of:
  - memory, intelligence, pattern recognition, ...

History of NNs - 2

Warren S. McCulloch and Walter Pitts, "A logical calculus of the ideas immanent in nervous activity", *Bulletin of Mathematical Biophysics*, 5:115-133, 1943.

- Historically very significant as an attempt to understand what the nervous system might actually be doing
- First to treat the brain as a computational organ
- Showed that their nets of "all-or-nothing" nodes could be described by propositional logic

History of NNs - 3

Donald O. Hebb, *The Organization of Behavior*, John Wiley & Sons, New York, 1949.

- **Hebb proposed a learning rule for NNs:**

"Let us assume then that the persistence of repetition of a reverberatory activity (or trace) tends to induce lasting cellular changes that add to its stability. The assumption can be precisely stated as follows: When an axon of cell A is near enough to excite cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place on one or both cells so that A's efficiency as one of the cells firing B is increased."

History of NNs - 4

Frank Rosenblatt, "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain", *Psychological Review*, 65:386-408, 1958.

- **used random, weighted connections between layers of nodes**
- **connection weights were updated according a a Hebbian-like rule**
- **was able to discriminate between some classes of patterns**

History of NNs - 5

Marvin Minsky and Seymour Papert, *Perceptrons, An Introduction to Computational Geometry*, MIT Press, Cambridge, MA, 1969.

- **AI community felt that NN researchers were overselling the capabilities of their models**
- **highlighted the theoretical limitations of the Perceptron at the time (which had been improved since the original version). Classic example is the inability to solve the XOR problem**
- **Effectively stopped NN research for many years**

History of NNs - 6

- **Some research continued:**
  - **Associative memories**
    - **James A. Anderson, " A Simple Neural Network Generating an Interactive Memory", *Mathematical Biosciences* 14:197-220, 1972.**
    - **Teuvo Kohonen, " Correlation Matrix Memories", *IEEE Transaction on Computers*, C-21:353-359, 1972.**
  - **Cognitron - the first multilayer NN**
    - **K. Fukushima, "Cognitron: A Self-organizing Multilayered Neural Network", *Biological Cybernetics*, 20:121-136, 1975.**
  - **Hopfield Networks**
    - **J. J. Hopfield, " Neural Networks and Physical Systems with Emergent Collective Computational Abilities", *Proceedings of the National Academy of Sciences*, 79:2554-2558, 1982.**

History of NNs – 7

- The limitations pointed out by Minsky and Papert were due the the fact that the Perceptron had only two layers (and was thus restricted to classifying linearly separable patterns)
- Extending successful learning techniques to multilayer networks was the challenge
- In 1986, several groups came up with essentially the same algorithm, which became known as *back-propagation*
- This led to the revival of NN research

History of NNs - 8

David D. Rumelhart, Geoffrey E. Hinton and Ronald J. Williams, "Learning Representations by Back-Propagating Errors", *Nature* 323:533-536, 1986.

- The idea of back-propagation is to calculate the *error* at the output layer, and then to trace the contributions to this error back through the network to the input layer, adjusting weights as one goes so as to reduce this error

History of NNs - 9

- Mathematically, this is a *gradient descent* training procedure
- In fact, back-propagation is the neural analogue of a gradient descent algorithm discovered earlier
  - Paul Werbos, "Beyond regression: New Tools for Prediction and Analysis in the Behavioral Sciences", Doctoral thesis, Harvard University, 1974.
- The back-propagation algorithm uses the *Chain Rule* from calculus to extend more traditional regression to multilayer networks

History of NNs - 10

- Probably the most common type of NN used to today is a multilayer feedforward network trained using back-propagation (BP)
- Often called a *Multilayer Perceptron* (MLP)
- Despite the title of Werbos' thesis, back-prop is now seen as a form of regression: a training set of input-output pairs is provided, and gradient descent is used to determine the the parameters of a model (the NN) to fit this training data

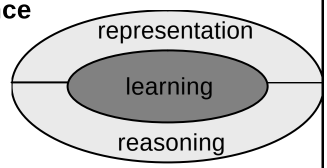
History of NNs - 11

- Other NN models have been developed during the last twenty years:
  - Adaptive Resonance Theory (ART)
    - pattern recognition networks where activity flows back and forth between layers, and "resonances" form
    - Gail Carpenter and Stephen Grossberg, "A Massively Parallel Architecture for a Self-Organizing Neural Pattern Recognition Machine", *Computer Vision, Graphics and Image Processing* 37:54, 1987.
  - Self-Organizing Maps (SOMs)
    - Also biologically inspired: *"How should the neurons organize their connectivity to optimize the spatial distribution of their responses within the layer?"*
    - Can be used for clustering (more next week)
    - Teuvo Kohonen, "Self-organized formation of topologically correct feature maps", Biological Cybernetics 43:59-69, 1982.

- Today, Neural Networks should be seen as part of a larger field sometimes called
  - *Soft Computing*
  - *Natural Computing*
  - *Artificial Intelligence*
  - *Machine learning*

AI and NN

- Definition of AI; Goal of AI
  - art of creating machine that performs tasks that requires intelligence when performed by people
  - study of mental faculties through the use of computational models
  - to make computers to perceive, reason and act
  - to develop machine that perform cognitive tasks
- functions of AI system
  - store knowledge
  - apply the knowledge to solve problems
  - acquire new knowledge thru experience
- Key components of AI
  - representation
  - reasoning
  - learning



AI

- AI is goal, objective, dream
- NN is a model of intelligent system
  - it is not the only system
  - Intelligent system is not necessarily same as human
    - Example : Chess machine
- Symbolic AI is a tool, paradigm toward AI
- NN can be a good tool toward AI

Representation in Symbolic AI

- By symbol structure
  - symbolic AI
  - well suited for human-machine communication
- Knowledge is just data
  - declarative
    - represented by facts
    - general inference procedure is used to manipulate the facts
    - PROLOG
  - procedural
    - knowledge is embodied in executable code
    - LISP

Lecture1-51

Reasoning

- Reasoning is ability to solve problem
  - must able to express and solve broad range of problems
  - must able to make explicit and implicit information known to it
  - must have control mechanism to select operators for a situation
- Problem solving is a searching problem
- deal with incompleteness, inexactness, uncertainty
  - probabilistic reasoning, plausible reasoning, fuzzy reasoning

Lecture1-52

Learning

- Model of Machine Learning (Figure 1.25)
  - Environment,
  - Learning element,
  - Knowledge base, and
  - performance cycle
- Inductive learning
  - generate rules from raw data
  - similarity-based learning, case-based reasoning
- deductive learning
  - general rules are used to determine specific facts
  - theorem proving
- Augmenting knowledge-base is not a simple task

Symbolic AI vs. NN

Level of explanation	■ Symbolic representation ■ cognition as sequential processing of symbolic representation	■ Parallel distributed processing ■ neurobiological explanation
Processing style	■ sequential, step-by-step ◆ logical-inference-like ◆ von Neumann machine	■ parallelism ◆ processing power in certain tasks such as sensory proc. ◆ flexibility ◆ robustness
Representation structure	■ quasi-linguistic ■ composition of symbolic expression ■ top-down	■ no structure ■ bottom-up

Symbolic AI + NN

- Structured connectionist model
- hybrid systems integrate them together
  - from NN
    - adaptivity
    - robustness
    - uniformity
  - from symbolic AI
    - representation
    - inference
    - universality
- Extracting rule from trained NN
    - if possible, it will be good for explanation

Advantage of Extracting rules from NN

- To validate NN by making internal state of NN accessible and understandable to users
- to identify regions of input space need more samples
- indicate the circumstances where NN fails to generalize
- to discover salient features for data exploration
- to traverse the boundaries of symbolic and connectionist approaches
- to satisfy the critical need for safety