Introduction To Artificial Neural Networks

Prof. Marcello Pelillo

Department of Computer Science
Ca' Foscari University, Venice

http://www.unive.it
http://www.dsi.unive.it

DARPA Neural Network Study

- "Over the history of computing science, two advances have matured: High speed numerical processing and knowledge processing (Artificial Intelligence). Neural networks seem to offer the next necessary ingredient for intelligent machines-namely, knowledge formation and organization."

- Neural network is a computational structure modeled after biological processes

- Neural network architectures are very different from traditional single processor computers

  1. Neural networks are adaptive, or trainable with data
     They are said to improve with experience; during training,
     neural networks update the interconnection weights

  2. Neural networks are naturally massively parallel, high speed,
     and fault tolerant (Information is stored in a distributed fashion);
     more interconnections than the processing units; processing
     power of a neural network ≡ the number of interconnection updates
     per second

- Artificial Intelligence versus neural network models
  Some have envisioned ANN as an alternative to AI, others believe this new-found faith in ANN to be naive

HISTORY

Early work (1940-1960)

- McCulloch & Pitts (Boolean logic)
- Rosenblatt (Learning)
- Hebb (Learning)

Transition (1960-1980)

- Widrow – Hoff (LMS rule)
- Anderson (Associative memories)
- Amari

Resurgence (1980-)

- Hopfield (Ass. mem. / Optimization)
- Rumelhart (Back-prop)
- Kohonen (Self-organizing maps)
- Hinton , Sejnowski (Boltzmann machine)

A Few Figures

Human cerebral cortex is composed of about 100 billion (10^{11}) neurons (nerve cells) of many different types.

Each neuron is connected to other 1000 / 10000 neurons, wich yields

10^{15} / 10^{14} connections

The cortex covers about 0.15 m² and is 2 – 5 mm thick

Biological Neural Networks

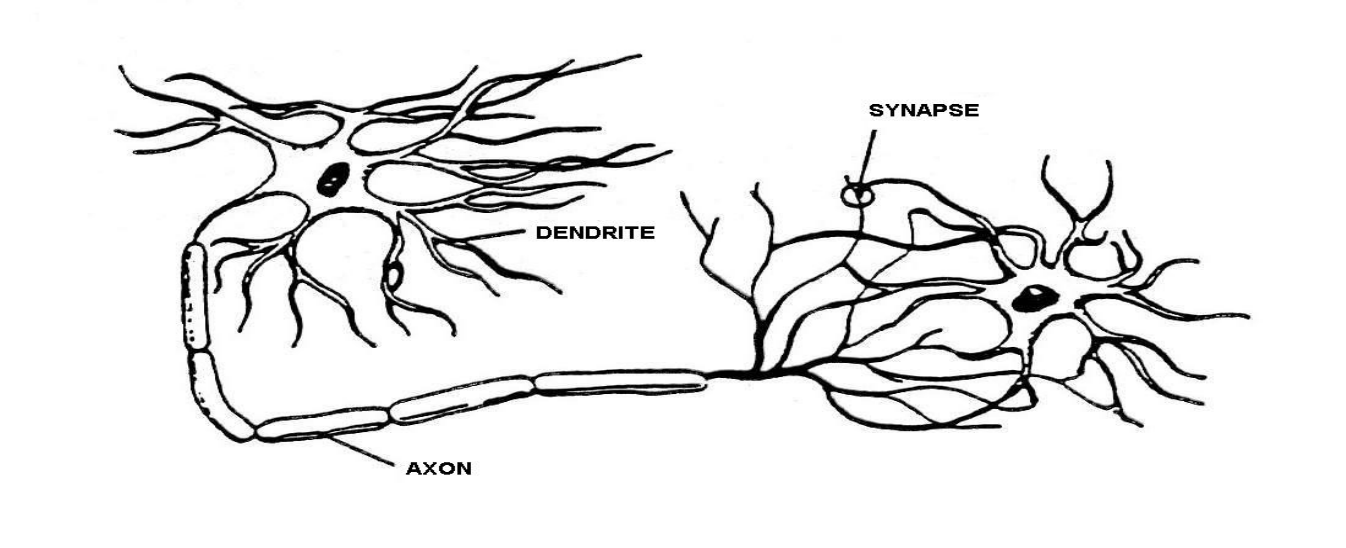
- Human cerebral cortex is composed of about 100 billion (10^{11}) neurons (nerve cells) of many different types; each neuron has roughly 1.000 dendrites (large fanout) that form ~100.000 bilion (10^{14}) synapses to other neurons; the system operates at 100 Hz or at some 10.000 (10^{16}) interconnections/sec., weighs ~3lbs, covers about 0.15 sq. m., and is about 2mm thick.

- Transmission of a signal from one cell to another at a synapse is a complex chemical process in which transmitter substances are released from the sending side of the junction which raises or lowers the electric potential of the receiving cell; if this potential exceeds a threshold, a pulse or action potential of fixed strength or duration is sent down the axon (cell has “fired”); after firing, the cell has to wait for a time called the refractory period before it can fire again.

- Processing in brain is broken down into groups or assemblies of neurons that act as functional units

Biological Neural System

- Generic neural structure
  - **Cell Body (Soma):** 5-10 microns in diameter
  - **Axon:** Output mechanism for a neuron; one axon/cell, but thousands of branches and cells possible for a single axon
  - **Dendrites:** Receive incoming signals from other nerve axons via synapse



Neural Dynamics

The transmission of signal in the cerebral cortex is a complex process:

electrical $\rightarrow$ chemical $\rightarrow$ electrical

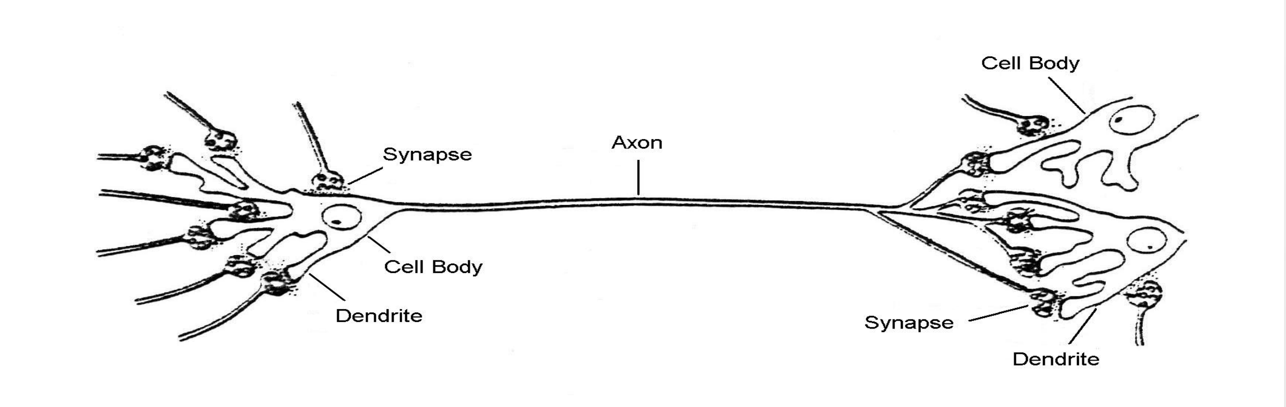
Simplifying :

1) The cellular body performs a “weighted sum” of the incoming signals

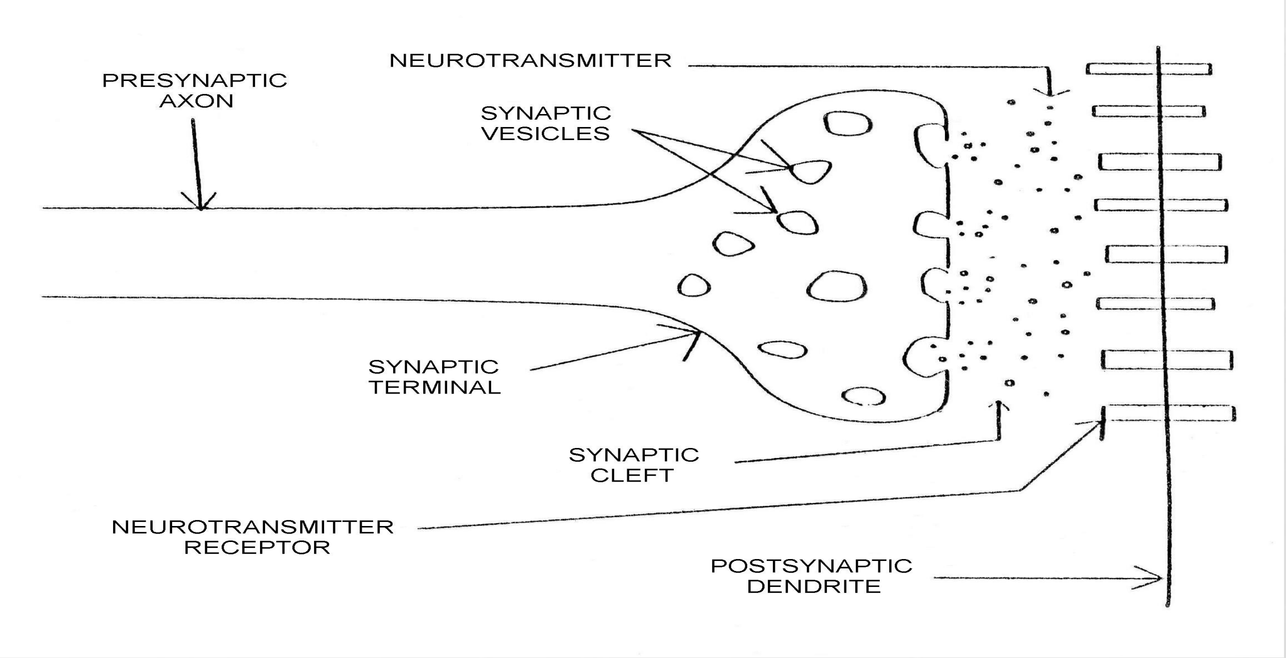
2) If the result exceeds a certain threshold value, then it produces an “action potential” which is sent down the axon (cell has “fired”), otherwise it remains in a rest state

3) When the electrical signal reaches the synapse, it allows the “neuro-transmitter” to be released and these combine with the “receptors” in the post-synaptic membrane

4) The post-synaptic receptors provoke the diffusion of an electrical signal in the post-synaptic neuron

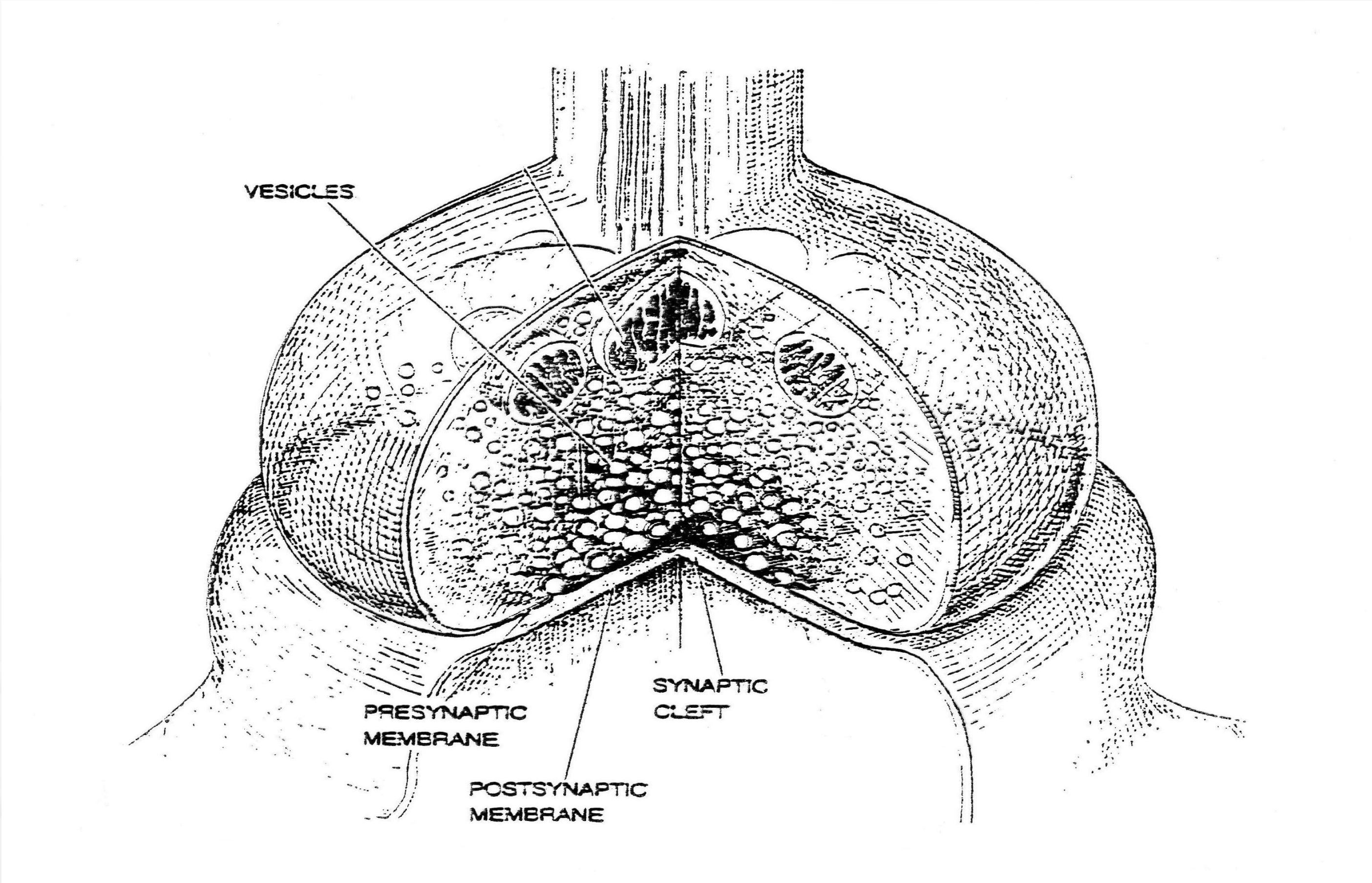


Biological Neurons



Synapse

SYNAPSES



SYNAPSE is the relay point where information is conveyed by chemical transmitters from neuron to neuron. A synapse consists of two parts: the knowblike tip of an axon terminal and the receptor region on the surface of another neuron. The membranes are separated by a synaptic cleft some 200 nanometers across. Molecules of chemical transmitter, stored in vesicles in the axon terminal, are released into the cleft by arriving nerve impulses. Transmitter changes electrical state of the receiving neuron, making it either more likely or less likely to fire an impulse.

Synaptic Efficacy

It's the amount of current that enters into the post-synaptic neuron, compared to the action potential of the pre-synaptic neuron.

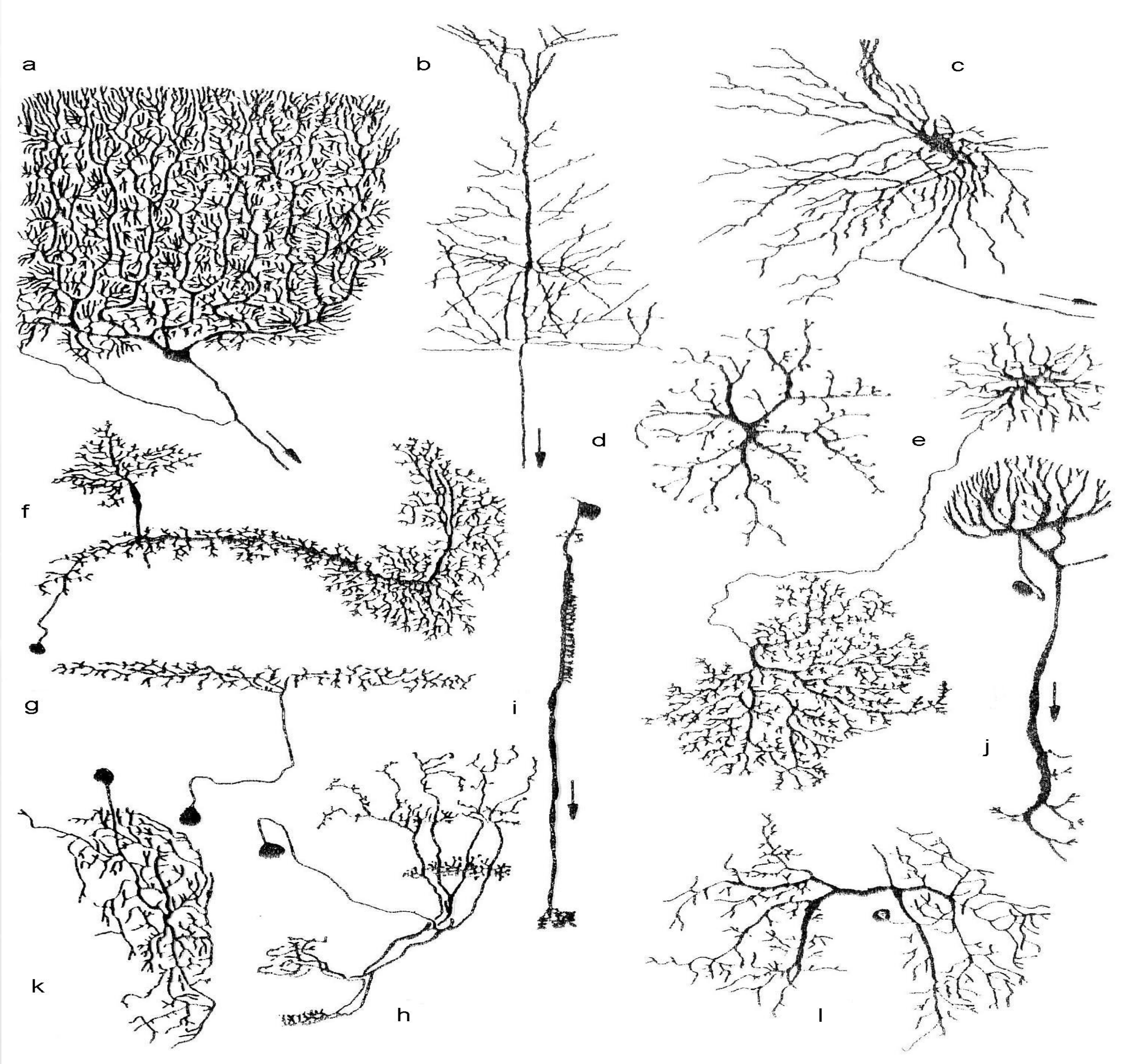
Learning takes place by modifying the synaptic efficacy.

Two types of synapses:

- **Excitatory** : favor the generation of action potential in the post-synaptic neuron
- **Inhibitory** : hinder the generation of action potential

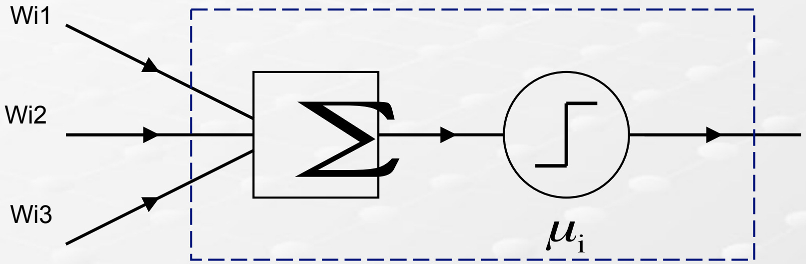
Anatomical Multiplicity Of Neurons

a Cellula di Purkinje (uomo)
b Cellula piramidale (coniglio)
c Motoneurone (gatto)
d Cellula orizzontale (gatto)
e Cellula orizzontale (gatto)
f Interneurone premotorio (cavalletta)
g Cellula visiva amacrina (mosca)
h Neurone multipolare (mosca)
i Neurone monopolare visivo (mosca)
j Interneurone visivo (cavalletta)
k Interneurone premotorio (gambero di fiume)
l Interneurone meccanosensoriale (gambero di fiume)



McCulloch and Pitts Model*

Neuron is modeled as a binary threshold unit



Schematic diagram of a McCulloch-Pitts neuron

The unit fires if the weighted sum $\sum_j w_{ij} n_j$ of the input reaches or exceeds the threshold μ_i

$$n_i(t+1) = \theta\ (\sum_j w_{ij} n_j(t) - \mu_i)$$

If neuron is firing, then n_i is 1, otherwise it is 0.

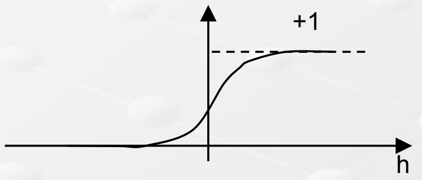
$\theta(x)$ is the unit step function

Wheights w_{ij} represent the strength of the synapse between neuron j and neuron i

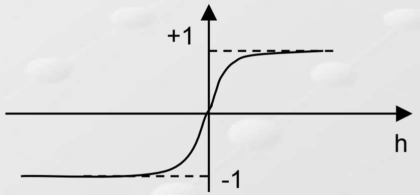
Continuous-Valued Units

- $g(h)$ is a continuous and differential function of h, e.g. linear and sigmoid. With such a $g(h)$, we can construct an error measure or cost function and then find a set of weights to minimize the cost function

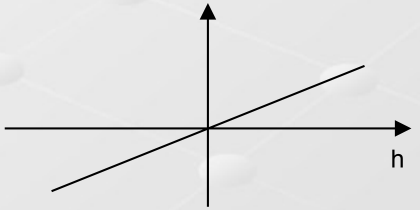
- Gradient descent learning is commonly used



Sigmoid unit



tanh(h) unit

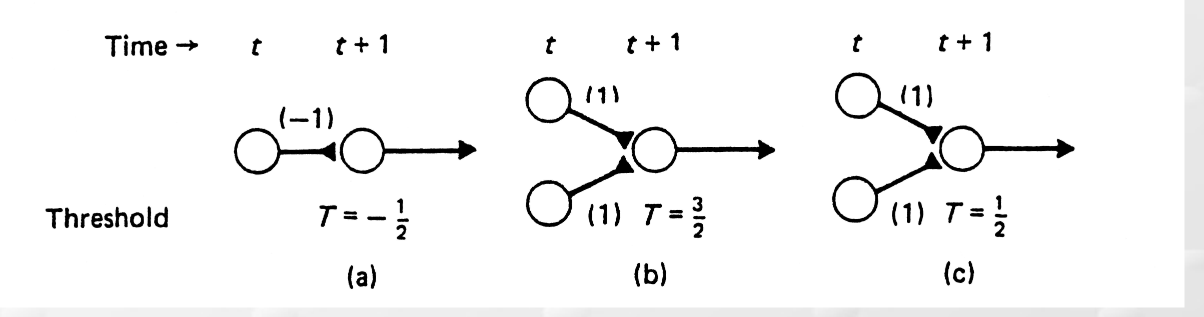


Linear unit

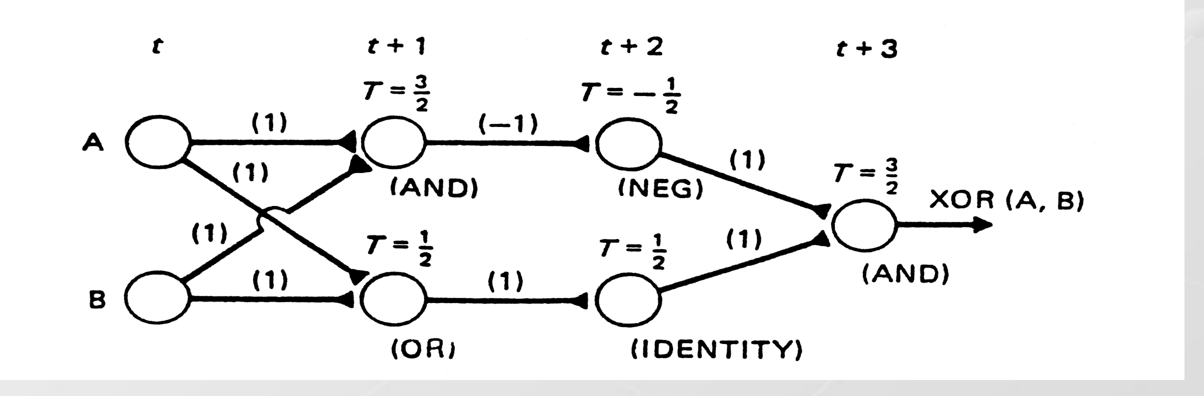
Proprietà del Modello M&P

Combinando opportunamente i neuroni di M&P si possono costruire reti in grado di realizzare qualsiasi operazione del calcolo proposizionale

Idea nuova: *il tempo*



Three elementary logical operations (a) **negation**, (b) **and**, (c) **or**. In each diagram the states of the neurons on the left are at time *t* and those on the right at time *t* +1.



The construction for the **exclusive or**

Testi di riferimento

J. Hertz, A. Krogh, R. Palmer
Introduction to the Theory of Neural Computation
Addison-Wesley

S. Haykin
Neural Networks – A Comprehensive Introduction
Prentice-Hall