

# Le reti neurali

Alberto Borghese

Università degli Studi di Milano  
Laboratory of Applied Intelligent Systems (AIS-Lab)  
Dipartimento di Informatica  
[alberto.borghese@unimi.it](mailto:alberto.borghese@unimi.it)



A.A. 2018-2019

1/41

<http://borghese.di.unimi.it>



## Sommario



Apprendimento supervisionato

I neuroni

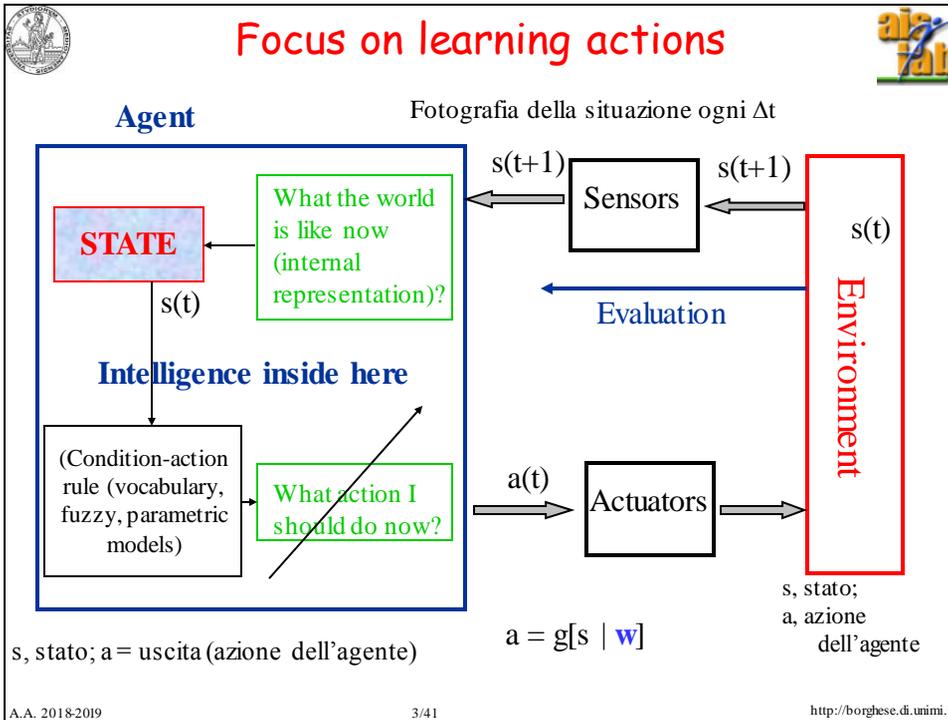
Le reti neurali

Apprendimento sociale

A.A. 2018-2019

2/41

<http://borghese.di.unimi.it>



**I vari tipi di apprendimento**  
**Apprendimento supervisionato**

$x(t+1) = f[x(t), a(t)]$       Ambiente  
 $a(t) = g[x(t)]$                       Agente

**Supervisionato** (learning with a teacher). Viene specificato per ogni pattern di input, il pattern desiderato in output.

Viene confrontato il valore prodotto dall'agente,  $a(s_t)$ , con  $a_{des}(s_t)$  desiderato, la differenza viene utilizzata per correggere la funzione  $g(\cdot)$ .

**a) Regressione:** a partire da un insieme di dati costruisco un modello (funzione  $g(\cdot)$ ) che predica in modo robusto  $a(s_t)$  per valori di  $s_t$  non ancora esaminati.

**b) Classificazione:** la funzione  $g(\cdot)$   $s_t$  a una tra un certo numero di classi predefinite.

A.A. 2018-2019 4/41 http://borgnese.di.unimi.it



## Ruolo del teacher



$$x(t+1) = f[x(t), a(t)] \quad \text{Ambiente}$$
$$a(t) = g[x(t)] \quad \text{Agente}$$

**Supervisionato** (learning with a **teacher**). Viene specificato per ogni pattern di input, il pattern desiderato in output.

Il teacher fornisce degli esempi: coppie di valori:  $\{x_t, a_{des}(x_t)\}$ .

$x_t$  e  $a_t$  possono essere conosciuti con un'approssimazione (errore, conoscenza incompleta, distribuzione stocastica...).

Il teacher non conosce tutte le possibili coppie:  $\{x_t, a(x_t)\}$ .



## Models



$$a = g[s \mid \mathbf{w}]$$

Which is the adequate value of  $\{w\}$ ?

How many  $\{w\}$ ?

How is the shape of  $g(\cdot)$  modified by the  $\{w\}$ ?

Which  $g(\cdot)$  shall we use?

=> Parametric models: combination of "simple" functions to obtain  $g(\cdot)$ .

One possibility is to derive such simple functions looking at the behaviour of human neurons.



## Sommario



Apprendimento supervisionato

**I neuroni**

Le reti neurali

Apprendimento sociale



## Brains cause minds (J. Searle)



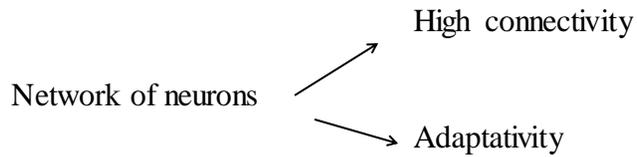
$10^{11}$  Neuroni interconnessi per:

Vedere, parlare, muoverci.....

Giocare a scacchi, appassionarci, discutere, imparare.....



## What is a brain?



Networks implement simple functions that are assembled dynamically according to the complex function required.

Spiking networks

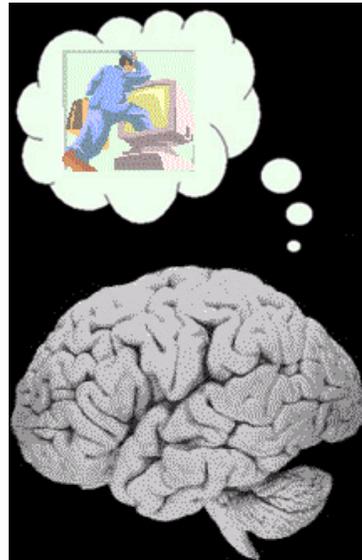


## L'intelligenza biologica



**Connessionismo cellulare** (K. Wernicke and R. Cajal, fine 1800)

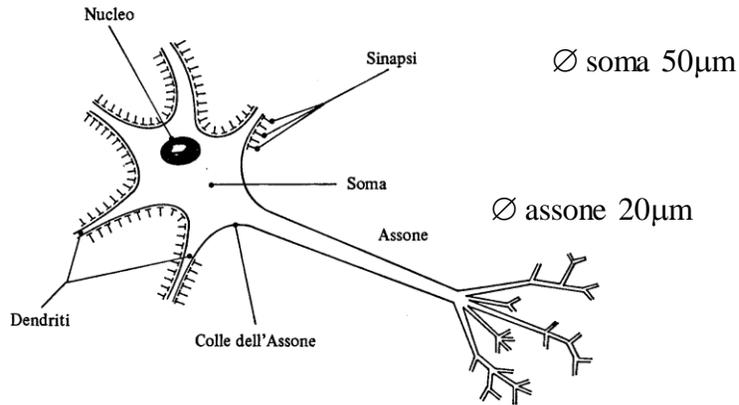
- I neuroni sono connessi tra loro in gruppi funzionali.
- Le connessioni sono in numero definito.
- Gruppi funzionali diversi danno origine a funzioni intellettive diverse.





## La struttura del neurone

Morfologicamente molto diversi, funzionalmente simili.



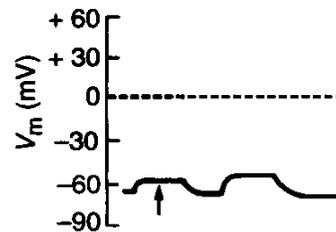
Dendriti: molti – input (da altri neuroni o recettori)

Assone: singolo, si diparte dal colle dell'assone – output (verso altri neuroni o effettori)

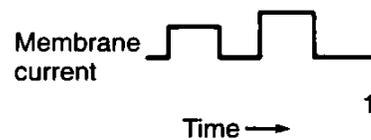


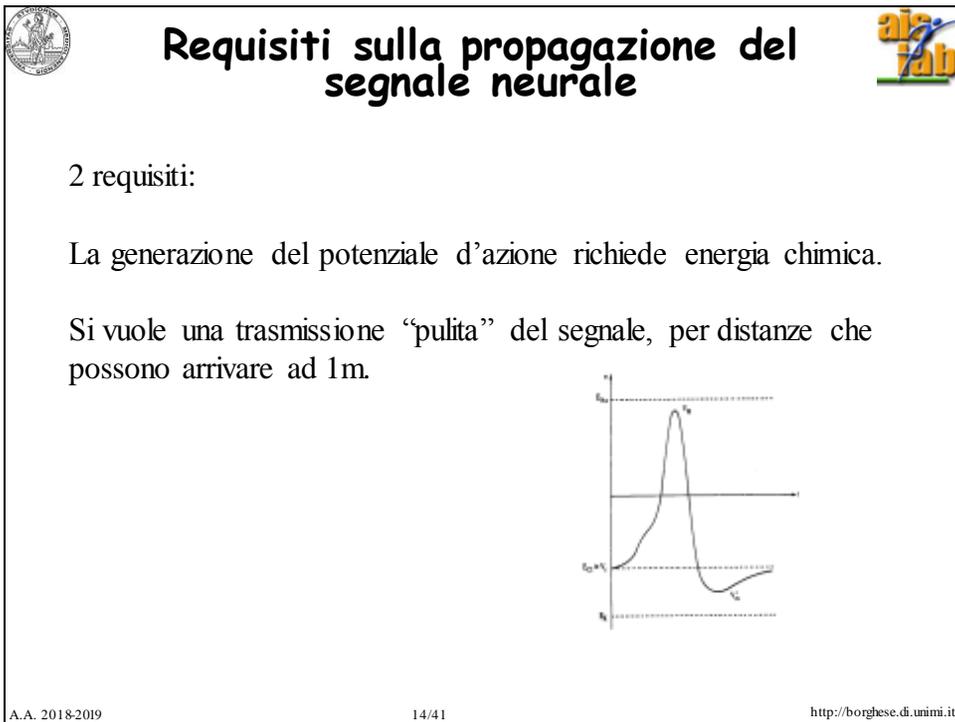
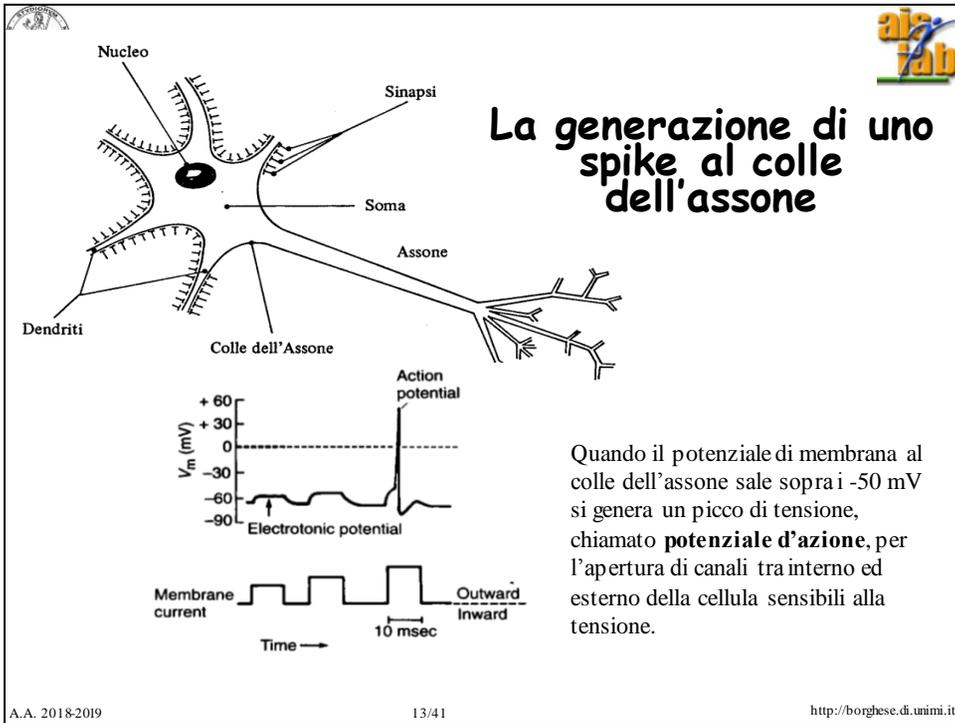
## Il Funzionamento sottosoglia

La tensione tra interno ed esterno della membrana cellulare varia seguendo la corrente che arriva dai dendriti.



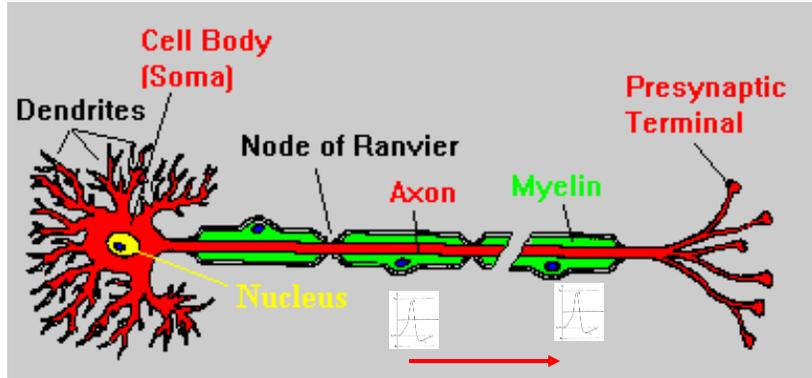
Potenziale di membrana







## Propagazione del segnale neurale



$\varnothing = 0.2 \div 20 \mu\text{m}$

**Mielina.** Protezione contro la dispersione del picco di tensione (filtraggio limitato).  
La demielinizzazione porta a una cattiva conduzione del segnale elettrico (spike)

Ripetizione del segnale ad ogni Nodo di Ranvier.

A.A. 2018-2019

15/41

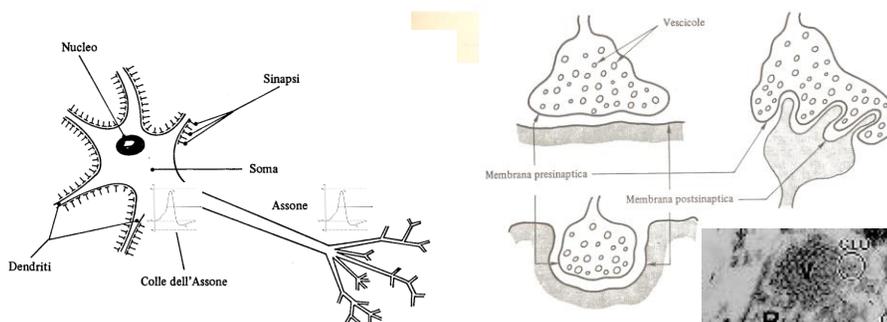
<http://borgnese.di.unimi.it>



## Le sinapsi



Sono l'interfaccia, la porta di I/O del neurone.



- Lo spike genera il rilascio di mediatori chimici.
- I mediatori chimici provocano una variazione di potenziale nella membrana post-sinaptica (variazione continua).
- Tempo di propagazione finito (0.5-1ms).

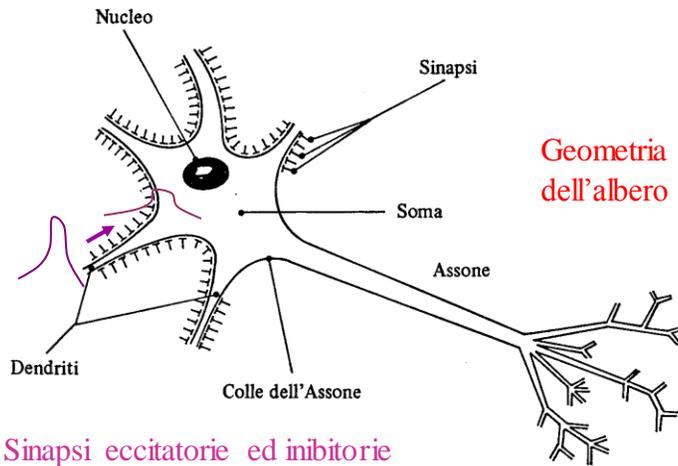
A.A. 2018-2019

16/41

<http://borgnese.di.unimi.it>



## L'integrazione nel corpo cellulare



Geometria computazionale dell'albero dendritico.

Sinapsi eccitatorie ed inibitorie  
Cambia la forma del segnale post-sinaptico a seconda della distanza dalla sinapsi.

Integrazione spazio-temporale degli input sinaptici.

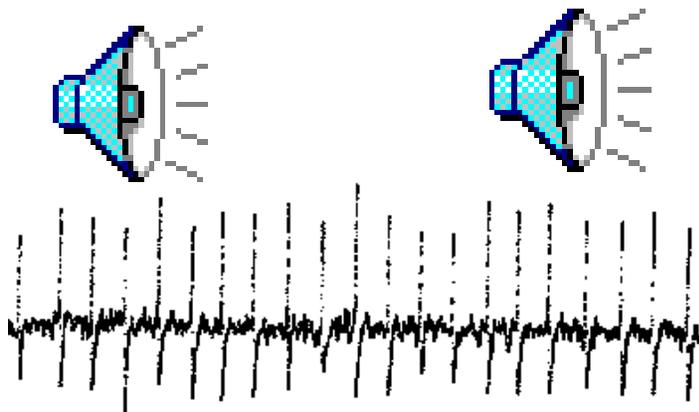
A.A. 2018-2019

17/41

<http://borgnese.di.unimi.it>



## Il suono del neurone



- Codice di frequenza.
- Periodo di refrattarietà.

A.A. 2018-2019

18/41

<http://borgnese.di.unimi.it>



## Brain characteristics



Number of neurons (adult)*	20,000,000,000 - 50,000,000,000
Number neurons in cerebral cortex (adult)	about 20,000,000,000 (some sources have incorrect number 8,000,000)
Number of synapses (adult)	$10^{14}$ (2,000-5,000 per neuron)
Weight Birth	0.3 kg, 1 y/o 1 kg, puberty 1.3 kg, adult 1.5 kg
Power consumption (adult)	20-40 Watts (0.5-4 nW/neuron)
Percentage of body	2% weight, 0.04-0.07% cells, 20-44% power consumption
Genetic code influence	1 bit per 10,000-1,000,000 synapses
Atrophy/death of neurons	50,000 per day (between ages 20 and 75)
Sleep requirement (adult)	average 7.5 hours/day or 31%
Normal operating temperature	$37 \pm 2^\circ\text{C}$
Maximum firing frequency of neuron	250-2,000 Hz (0.5-4 ms intervals)
Signal propagation speed inside axon	90 m/s sheathed, <0.1 m/s unsheathed
Processing of complex stimuli	0.5s or 100-1,000 firings

From Vadim Gerasimov's slides



## Sommario



Apprendimento supervisionato

I neuroni

**Le reti neurali**

Apprendimento sociale



## Le reti neurali



Se il neurone biologico consente l'intelligenza, perché non dovrebbe consentire l'intelligenza artificiale un neurone sintetico?

“.. a neural network is a system composed of *many simple processing elements* operating in *parallel* whose function is determined by *network structure, connection strengths*, and the *processing performed at computing elements* or nodes. ... Neural network architectures are inspired by the architecture of biological nervous systems, which use many simple processing elements operating in parallel to obtain high computation rates”. (DARPA, 1988)....

Now, this is called learning with Kernels



## A cosa servono?



Le reti neurali offrono i seguenti specifici vantaggi nell'elaborazione dell'informazione:

- Apprendimento basato su esempi (non è richiesta l'elaborazione di un modello aderente alla realtà)
- Autoorganizzazione dell'informazione nella rete
- Robustezza ai guasti (codifica ridondante dell'informazione)
- Funzionamento in tempo reale (realizzazione HW)
- Basso consumo (0.5nW ÷ 4nW per neurone, 20W per il SN).





## Cosa sono le reti neurali artificiali?



- Le reti neurali sono modelli non lineari per l'**approssimazione** della soluzione di problemi dei quali non esiste un modello preciso (o se esiste è troppo oneroso computazionalmente). I parametri dei modelli risultanti (semiparametrici) vengono calcolati mediante l'utilizzo di esempi (dati di ingresso e uscita desiderata). Connessioni con il dominio della statistica.
- Vengono utilizzate soprattutto per la classificazione e la regressione.
- Sono un capitolo importante negli argomenti di intelligenza artificiale.
- Da un altro punto di vista possono essere utilizzate per lo studio delle reti neurali naturali, ovvero dei processi cognitivi.
- Sono state incorporate nel "machine learning".



## Il modello "rete neurale"



Modello: rappresentazione parametrizzata di un fenomeno

$$y = f(x | \mathbf{w})$$

Dove i parametri  $\mathbf{w}$  definiscono la forma di  $f$  e quindi della mappatura tra gli ingressi e le uscite.

I parametri  $\mathbf{w}$  vengono determinati per ottenere una mappatura  $f(\cdot)$  congruente con i dati in ingresso e uscita misurati (apprendimento supervisionato).

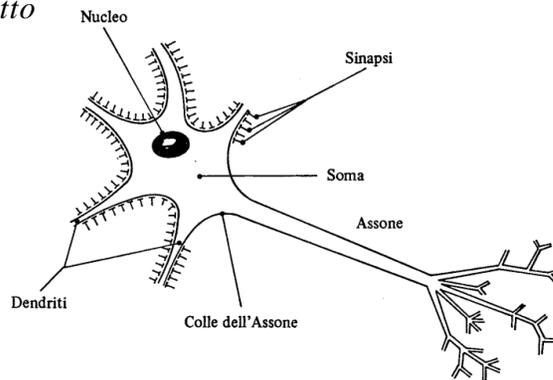
In particolare  $f(\cdot)$  è ottenuta combinando assieme tante funzioni elementary che operano trasformazioni stereotipate sui dati (e.g. Gaussiane, Sigmoidali, Lineari ...).



## Il neurone artificiale



- *Potenziale di azione (tutto o nulla).*
- *Integrazione nel soma.*
- *Soglia di attivazione.*



Neurone come elemento di calcolo universale: in grado di calcolare qualsiasi funzione logica (cioè implementabile in un computer).

A.A. 2018-2019

25/41

<http://borgnese.di.unimi.it>



## Il modello di McCulloch-Pitts



• La variazione delle forma d'onda del potenziale di membrana lungo il dendrita non viene considerata.

• Gli input non sono sincroni.

$$y_i(t+1) = \Theta(w_{ij}u_j(t) - \mu_i)$$

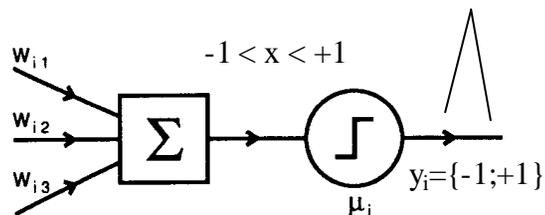
• Le interazioni tra input non sono lineari.

$$\Theta(x) = \begin{cases} 1 & \text{se } x \geq 0 \\ -1 & \text{altrimenti} \end{cases}$$

• I pesi sono supposti costanti.

• Descrizione nel tempo.

Sono state pensate per calcolare **funzioni logiche (V o F)**.



McCulloch-Pitts (1943)

A.A. 2018-2019

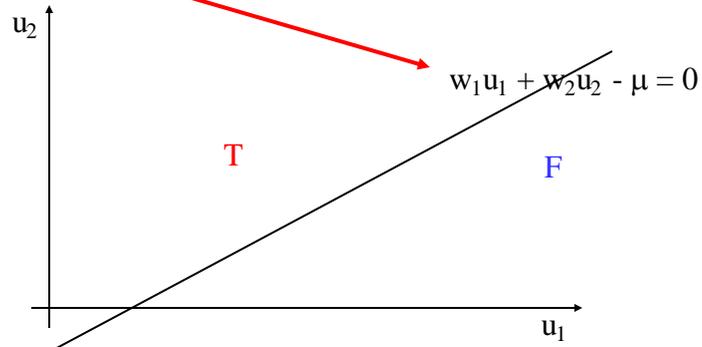
i.it



## Rappresentazione della retta



$$y_i(t+1) = \Theta (w_j u_j(t) - \mu_i)$$



La retta taglia in due il piano

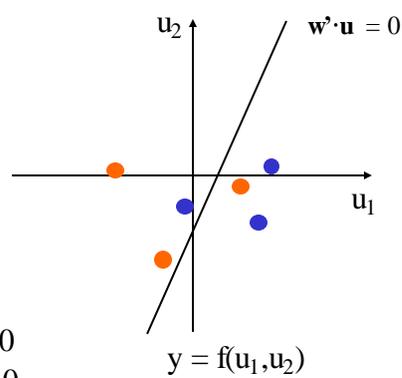
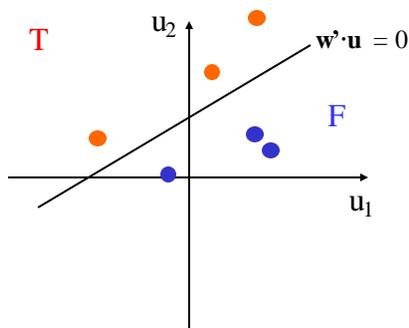


## Funzioni linearmente separabili (separabili mediante una retta)



Linearmente separabile

Non linearmente separabile



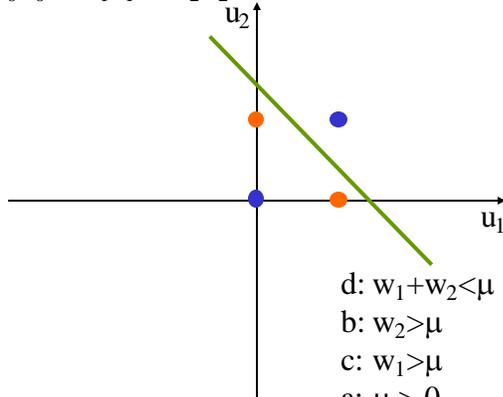
- $y > 0$
- $y < 0$



# La "morte" del neurone di McCulloch-Pitts (Minsky, 1969): XOR



$$w_0u_0 + w_1u_1 + w_2u_2 = 0$$



$u_1$	$u_2$	$y$
0	0	-1
0	1	1
1	0	1
1	1	-1

a  
b  
c  
d

- d:  $w_1 + w_2 < \mu$
- b:  $w_2 > \mu$
- c:  $w_1 > \mu$
- a:  $\mu > 0$

Il sistema di 4 equazioni non è risolvibile.

- $y(u_1, u_2, 1) = 1$
- $y(u_1, u_2, 1) = -1$

$w_1, w_2 > \mu$  e  $w_1 + w_2 < \mu$  Impossibile!!

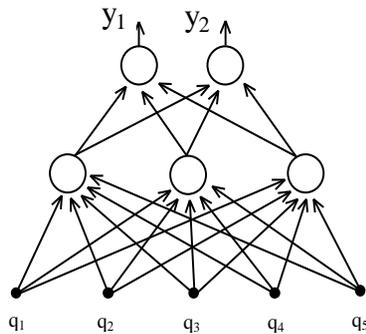
Si possono imparare solamente funzioni linearmente separabili

A.A. 2018-2

fi.unimi.it



# Spiking neurons



**Spiking neurons.** Sono neuroni la cui uscita è il singolo spike. Modellazione realistica (e.g. McCullochPitts). **Spike del neurone.** Registrazione dell'uscita di un insieme di neuroni nel tempo.

**Connessionismo classico.** Uscita compresa tra min - Max. **Frequenza di scarica.** Frequenza di scarica in uscita in funzione della frequenza di scarica dei vari ingressi.

A.A. 2018-2019

30/41

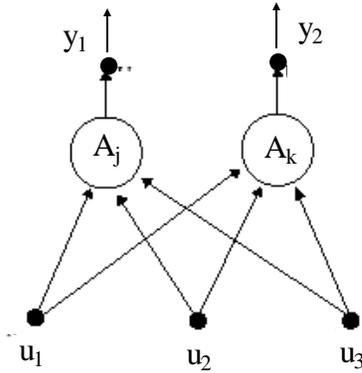
http://borghese.di.unimi.it



## La rete neurale ad un livello



La rete opera una trasformazione dallo spazio di input allo spazio di output. Visione “spaziale” e non temporale.



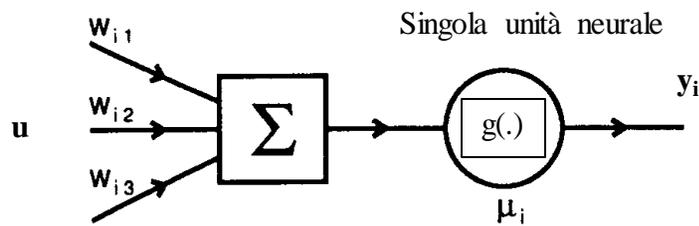
$$y_i = g(w_{ij}u_j - \mu_i)$$

La trasformazione o mappatura dipende dai parametri  $\{w_{ij}\}$  e  $\{\mu_i\}$  in modo tale che la rete neurale approssimi la trasformazione tra i pattern di input e di output.

Se  $g(\cdot) = 1$ , la rete diventa un modello lineare:  $y_i = w_{ij}u_j - \mu_i$

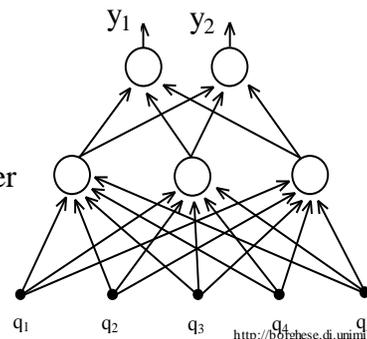


## Una rete neurale a più livelli



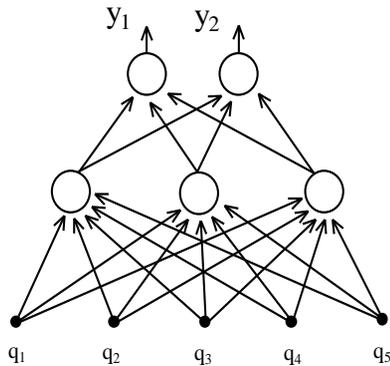
$$y_i = g(w_{ij}u_j - \mu_i)$$

Unità nascoste – Hidden layer





## Caratteristiche



Livelli di unità di attivazione

Collegamento in cascata

Input convergenti, output divergenti.

Capacità di approssimazione universale

**Perceptrone:** layered networks, flusso unidirezionale dell'elaborazione.

L'output viene interpretato come frequenza di scarica del neurone d'uscita della rete.

A.A. 2018-2019

33/41

<http://borgnese.di.unimi.it>



## Complessità della funzione realizzabile

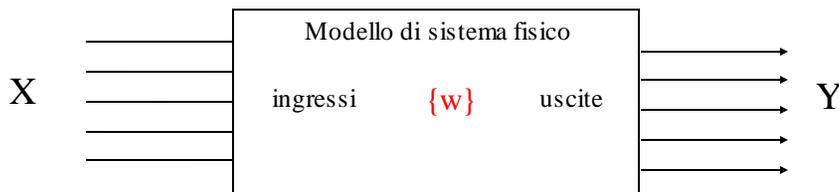


Quanti più neuroni artificiali vengono connessi tanto più la funzione complessiva approssimabile diviene più complessa

$$Y = |y_1, y_2, y_3, \dots, y_n|^T$$

$$y_i = g(X)$$

$$X = |x_1, x_2, x_3, \dots, x_m|^T$$



Reti neurali = approssimatori universali.

A.A. 2018-2019

34/41

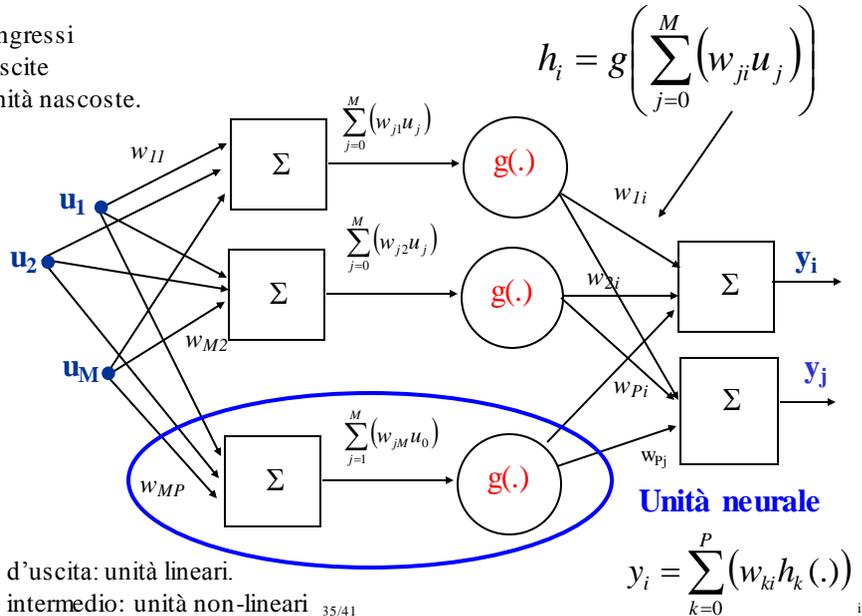
<http://borgnese.di.unimi.it>



## MLP : Multi-layer Perceptron



- M ingressi
- N uscite
- P unità nascoste.



Livello d'uscita: unità lineari.

Livello intermedio: unità non-lineari

35/41



## Riassunto - topologia



I neuroni connessioneisti sono basati su:

- Ricevere una somma pesata degli ingressi.
- Trasformarla secondo una funzione non-lineare (scalino o logistica)
- Inviare il risultato di questa funzione all'uscita o ad altre unità.

Le reti neurali sono topologie ottenute connettendo tra loro i neuroni in modo opportuno e riescono a calcolare funzioni molto complesse.



## Riassunto - Apprendimento



Algoritmi iterativi per adattare il valore dei parametri (pesi).

Definizione di una funzione costo che misura la differenza tra valore fornito e quello desiderato.

Algoritmo (gradiente) che consente di aggiornare i pesi in modo da minimizzare la funzione costo.

Training per pattern (specializzazione) o per epoche.



## Sommario



I neuroni

Le reti neurali

**Apprendimento sociale**



## Mirror neurons - neuroni specchio



<http://www.youtube.com/watch?v=O3-wegp1ovM>

I Neuroni specchio



A.A. 2018-2019

39/41

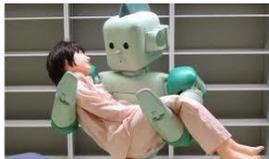
<http://borgese.di.unipi.it>



## Social robotics



«Per definizione un robot sociale dovrebbe comunicare ed interagire con gli umani o con altri esseri viventi in situazioni che possono essere definite cooperative. Ma anche comportamenti non cooperativi possono essere considerati sociali in alcune situazioni. Il robot può, ad esempio, esibire un comportamento competitivo all'interno del contesto di un gioco. Il robot potrebbe anche interagire, in alcuni casi con nessuna o con una minima comunicazione. Potrebbe ad esempio consegnare strumenti ad un astronauta che lavori su una stazione spaziale» (Wikipedia).



Video on Qrio dancing salsa

A.A. 2018-2019

40/41

<http://borgese.di.unipi.it>



## Sommario



I neuroni

Le reti neurali

Apprendimento sociale