

AUTOMI A GIUDIZIO



Nozioni storiche e concetti principali per un approccio alla neuroinformatica (giuridica)

a cura di Pier Giuseppe Milanese

*Università degli Studi di Pavia – Facoltà di Giurisprudenza – A. A. 2015/2016
Corso di informatica giuridica e logica giuridica – Prof. ROMANO ONEDA*

L'INTEGRAZIONE UOMO/MACCHINA - 1

PREMESSA

*E' indubbio che in futuro robot e softbot interagiranno sempre più strettamente con gli esseri umani, non solo condizionandone i comportamenti (e perciò «automatizzando» l'umano), ma addirittura guadagnando sempre più livelli di autonomia nel momento stesso in cui le macchine saranno in grado di operare sulla base di metodologie di apprendimento **non supervisionato**, ossia in modo sempre più autonomo e **induttivo**.*

Valutazione dell'apprendimento. Il problema, prima ancora che sul piano tecnico, può essere proposto in termini epistemologici, se non addirittura «pedagogici» (nel caso specifico si dovrebbe parlare di ... **automatagogici!**). La domanda è: *che cosa significa «corretto apprendimento»?* - prevedendo in tal modo che l'automa possa conseguire un «diploma» che ne attesti il buon livello di preparazione utile a svolgere particolari mansioni.

Il **PAC-learning** (Probably Approximately Correct learning) è una procedura matematica di analisi dell'apprendimento meccanico introdotto nel 1984 da L. Valiant e successivamente perfezionato in conseguenza dei livelli di complessità degli automi.

L'INTEGRAZIONE UOMO/MACCHINA - 2

PREMESSA

Questioni etico- giuridiche. Un soggetto può essere definito «responsabile» di x solo se ha il pieno controllo su x. L'attribuzione di responsabilità ad una macchina che apprende è una recente acquisizione del diritto all'interno di una più vasta problematica sulla responsabilità in cui la catena causale che provoca un danno non è chiaramente riconoscibile per cui nessun soggetto è palesemente identificabile come «colpevole»

Sul piano canonico, l'attribuzione di responsabilità che più si avvicina al caso degli automi sarebbe assimilabile alla «responsabilità del tutore» o del genitore nei confronti del minore o del possessore di animali per i danni che questi soggetti possano arrecare.

Tuttavia, trattandosi di macchine – e perciò in ultima istanza prodotti che vengono fabbricati e venduti – la catena di responsabilità diventa più complessa anche in nome del principio storico *ubi commoda, ibi incommoda* (chi ricava dei vantaggi, deve farsi carico anche degli svantaggi.) Si può parlare in questo caso di **responsabilità retrospettiva.**

Tuttavia, a seguito dell'incremento del grado di autonomia degli automi che loro di operare scelte elaborando una «propria esperienza» - anche sulla base di una capacità di calcolo superiore alla nostra – il discorso sulla «responsabilità» si fa estremamente più complesso.

AUTOMI - 1

Un automa è una macchina in grado di operare da sola oppure una macchina a meccanismo controllato, studiata per eseguire un numero determinato di operazioni in risposta ad istruzioni predeterminate. L'automata consiste di **stati** e **transizioni**. A fronte di un segno di input, l'automata compie una transizione ad un altro stato, secondo una funzione di transizione.

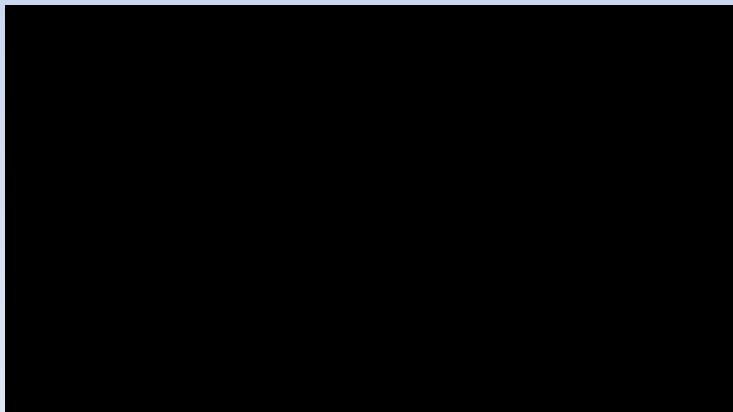
La teoria degli automi è strettamente connessa con la teoria dei linguaggi formali. Un automa è la rappresentazione finita di un linguaggio formale che può aspirare a infinite configurazioni. Per questo sono spesso classificati a seconda della classe di linguaggi che sono in grado di riconoscere.

*La caratteristica-base di queste macchine è che accettano un set **finito** di **input** e producono un set **finito** di **output** e possono presentare un determinato set di stati la cui definizione dipende dal tipo di automa.*



Automa a stati finiti. Un automa di questo tipo è di tipo *deterministico*, caratterizzato da stati che contengono un numero finito di elementi ; è in grado di gestire una quantità di memoria che non è però in grado di conservare. Un banale esempio può essere offerto da un distributore di bevande che riceve monetine (input) e può presentare 3 stati a) monete insufficienti b) monete sufficienti con distribuzione di bevanda c) monete eccedenti con distribuzione di bevande e restituzione della somma eccedente.

AUTOMI - 2



Macchina di Turing. Descrive un automatismo computazionale che sta alla base della architettura dei moderni computer. Essa è costituita da una testina *scorrevole nei due sensi* su un nastro costituito da un numero infinito di celle contenente simboli di un alfabeto (ad esempio binario) con cui la testina opera - in modalità lettura/scrittura/cancellazione/sostituzione. Ogni posizione della testina corrisponde ad uno «stato» della macchina associato alle istruzioni circa l'operazione che la testina deve eseguire su quella cella e allo spostamento verso un'altra cella (o stato)

Il gioco della vita. E' la versione più nota di un «automa cellulare». Rappresenta un implementazione della macchina di Turing. E' un gioco nel senso che è la macchina che gioca di sola **modificando i propri stati** a seconda della configurazione assunta dallo stato precedente. Le singole cellule nascono, muoiono, si conservano in base a dei parametri determinati dal rapporto con le cellule più vicine.

IL CONCETTO DI «APPRENDIMENTO»

Cosa significa «apprendere». Apprendere significa avere la possibilità di modificare le proprie risposte comportamentali a seguito della interazione con l'ambiente. Queste modificazioni, sulla base della esperienza, tendono infine a costruire o a consolidare **schemi generali di risposta** a determinati *set* di impulsi ambientali. E' un processo complesso, non riconducibile alla sola capacità di reagire ad uno stimolo o di conservare una esperienza. *«Apprendere» significa piuttosto «far tesoro» della passata esperienza per poter gestire l'esperienza stessa in modo tale da perseguire, in tempi sempre più rapidi, degli obiettivi prefissati.*

L'apprendimento «naturale». Gli organismi viventi sono in grado di modificare la loro struttura e di elaborare più sofisticate e ingegnose strategie atte a moltiplicare le opportunità adattative. Grazie allo sviluppo del sistema nervoso e alla sua centralizzazione (cervello) questa capacità di registrare, elaborare dati e rispondere in vista di obiettivi sempre più diversificati e complessi si è enormemente moltiplicata.

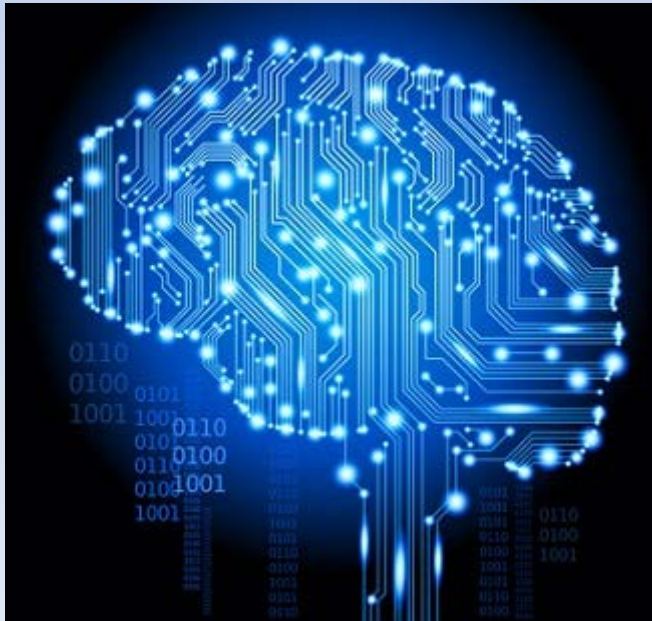
Macchine che «imparano». A. Samuel definì una **machine learning** come «un campo di studi che offre ai computer la capacità di imparare senza essere esplicitamente programmati». Questo non significa che la macchina non debba essere a sua volta «istruita». Vengono infatti approntati degli **algoritmi di apprendimento** che «insegnano» alla macchina come fare previsioni a partire da un *set* di dati e di optare per questa o quella soluzione - gestendo direttamente l'ingresso, la configurazione e il peso dei dati - al fine di ottenere un risultato che più si avvicini a quello desiderato.

IL MODELLO BIOLOGICO

IL NEURONE

In una struttura fortemente differenziata come quella cerebrale non esiste una sola tipologia valida per tutte le popolazioni neuronali. Si è perciò cercato di lavorare su un modello classico che meglio sintetizzi le caratteristiche specifiche di questa minuscola centralina in grado di ricevere, elaborare e trasmettere informazioni mediante impulsi elettrici e stimolazioni neurochimiche.

*Nel comune modo di pensare usiamo paragonare, anche se in modo improprio, il cervello ad un computer. In realtà **ogni singolo neurone** è un computer, per cui il cervello è piuttosto una rete costituita da 100 miliardi di computer che lavorano in sincronia e si aggiornano tra di loro.*



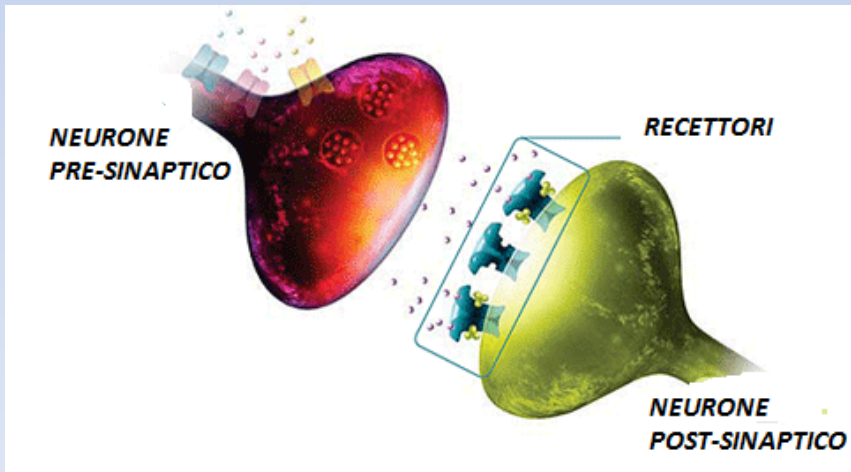
Ciascun neurone è connesso a migliaia e migliaia di altri neuroni dando origine ad un sistema di impressionante vastità, sensibilità ed elasticità, tale da non poter essere riprodotto dalle tecnologie disponibili. Da questo universo biologico possiamo solo estrapolare dei modelli semplificati applicabili in contesti finiti e con estensione e capacità operative limitate.

IL NEURONE - 1

Un neurone è in grado di ricevere, generare e trasmettere ad altri neuroni

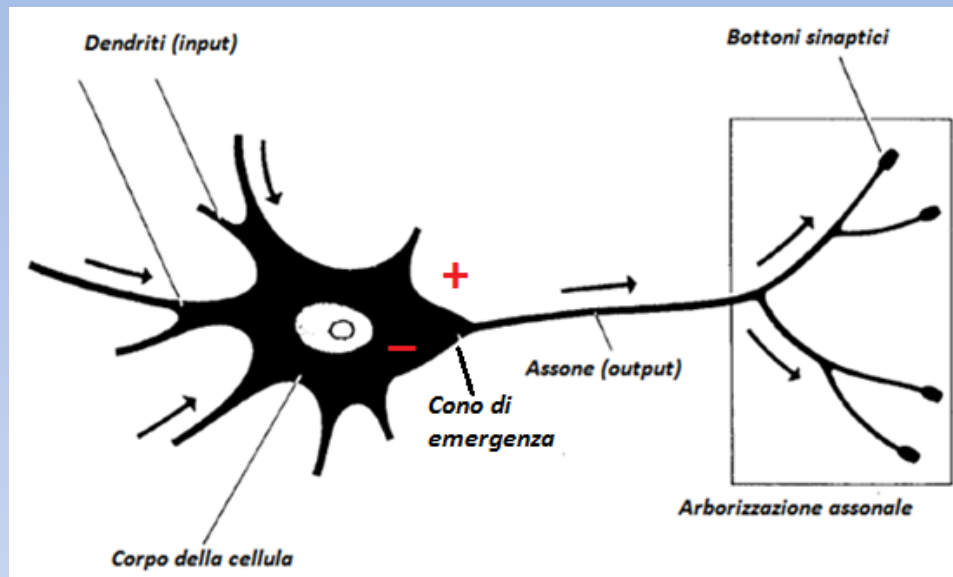
- a) Un segnale elettrico (sinapsi elettrica)
- b) Un segnale chimico (sinapsi chimica)

La **sinapsi elettrica** consente il **passaggio di corrente** da un neurone all'altro ed avviene attraverso particolari **giunzioni** (*gap junctions*). La trasmissione elettrica consente una attivazione simultanea di neuroni a prescindere da ogni altra forma di condizionamento. Le sinapsi elettriche consentono di gestire risposte veloci e sincroniche per particolari circuiti neuronali.



La **sinapsi chimica** è un processo **circoscritto e locale** con cui il neurone «deposita» un messaggio neurochimico nella «cassetta postale» dei neuroni riceventi. I messaggi sono sostanze chimiche *qualitativamente* differenziate, i **neurotrasmettitori**. Non se ne conosce il numero ma si ritiene siano molte decine. La «cassetta postale» che si apre, ricevendo, si chiama «recettore». Le svariate famiglie di neuroni sono specializzate a trasmettere o a ricevere solo alcuni tipi di messaggi. Insomma, come nelle buone famiglie, non si apre a chiunque ... suoni il campanello!

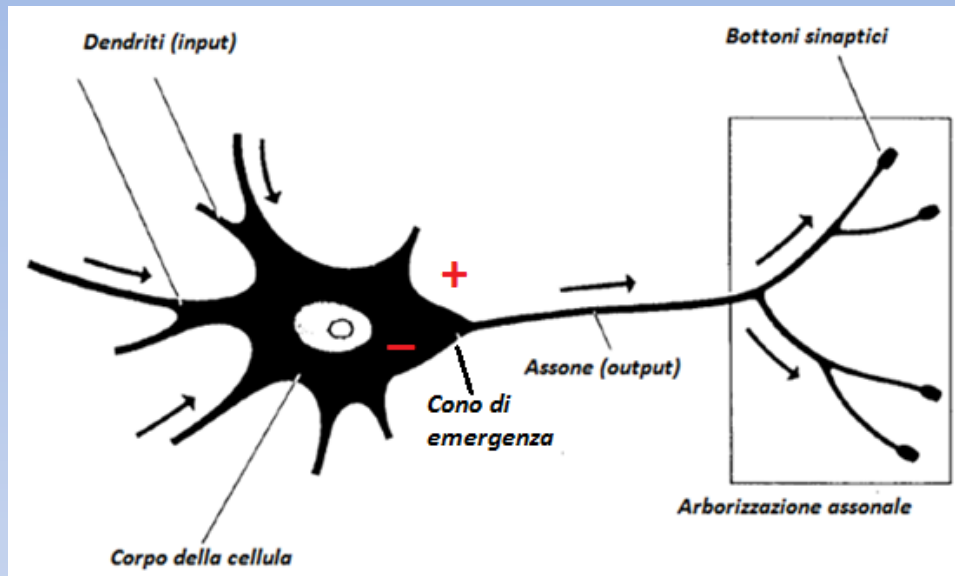
IL NEURONE - 2



Il neurone a riposo. Il corpo cellulare è avvolto da una **membrana** che separa l'ambiente interno/esterno e che è **carica elettricamente** a seguito della differenza di potenziale determinata dalla diversa concentrazione di ioni (sodio, potassio, calcio ecc.) tra l'interno della membrana e l'ambiente esterno.

Questa differenza fa sì che in situazione di riposo, l'interno della membrana abbia polarità negativa (-) con valore di -60 mVolt circa, mentre l'esterno/ambiente ha polarità positiva (+).

IL NEURONE - 3

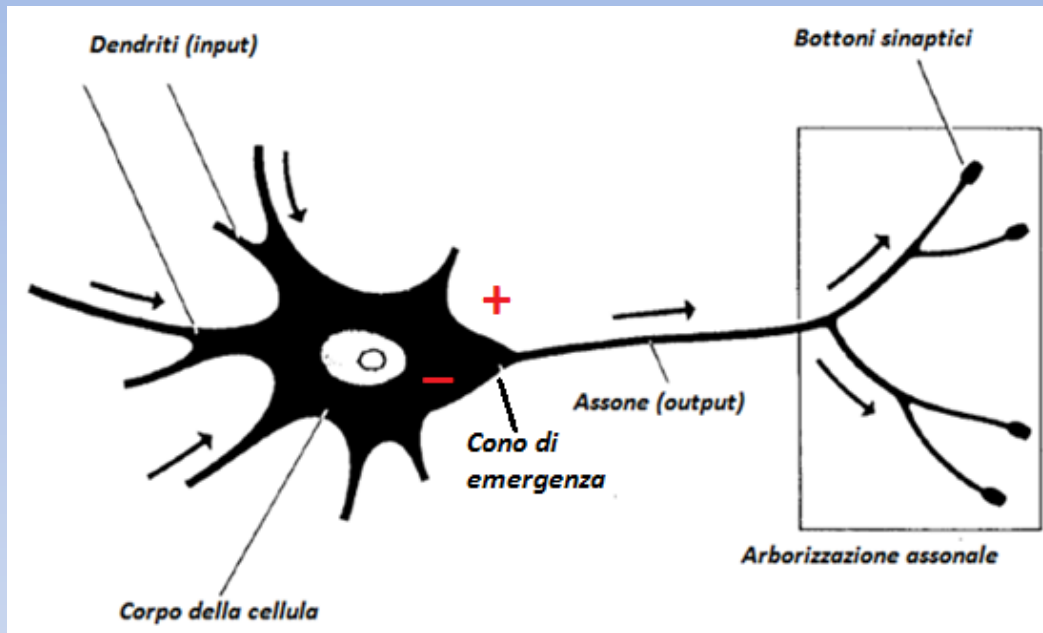


L'impulso nervoso. Se uno stimolo in entrata supera una certa **soglia** di intensità tale da provocare una leggera **depolarizzazione della membrana** (da -60 a -50 mV) si innesta una **reazione a catena**.

Lo stimolo apre le porte di ingresso per cui ioni positivi invadono l'interno determinando una **inversione di polarità** che può raggiungere un picco di $+50$ mV. L'apertura dura circa 1 ms. Poi le porte si chiudono ed inizia il processo inverso per cui gli ioni positivi vengono nuovamente espulsi. La polarità interna ridiscende di valore (fino a -70 mV) per poi raggiungere il suo valore di riposo.

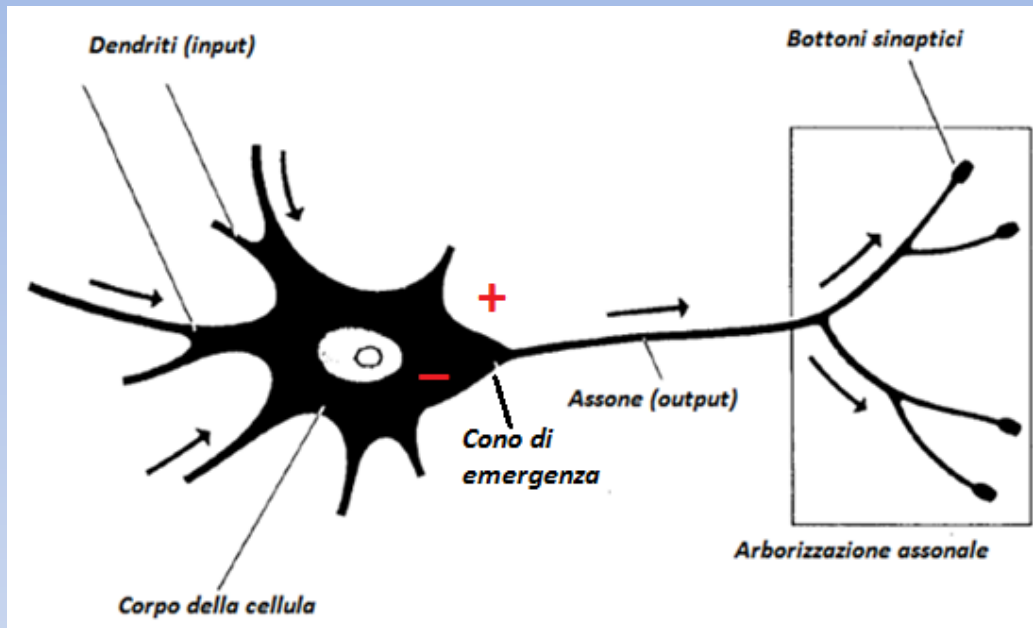
Questo comportamento genera **un'onda** che si propaga con ampiezza **costante**, andamento **unidirezionale** e **velocità** che può raggiungere i 150 m/s. Essa corrisponde a ciò che nel linguaggio comune chiamiamo **impulso nervoso**. Un impulso nervoso, nel caso più semplice, è quello stimolo che ad esempio è in grado di far contrarre un muscolo.

IL NEURONE - 4



Nel modello vediamo che l'assone termina con delle vescichette che contengono una sostanza chimica, il **neurotrasmettitore** che, al passaggio dell'impulso nervoso, viene rilasciato nel corpo della cellula ricevente dotata di **recettori** che si aprono legandosi al neurotrasmettitore. **Questa apertura libera risorse interne al neurone che ne determinano il comportamento.** Le sostanze attive che veicolano messaggi neurochimici, finora conosciute sono circa un centinaio, ma le più importanti sono una decina. Alcune sostanze (eccitatorie) «accendono» l'attività del neurone, altre (inibitorie) la spengono.

IL NEURONE - 5



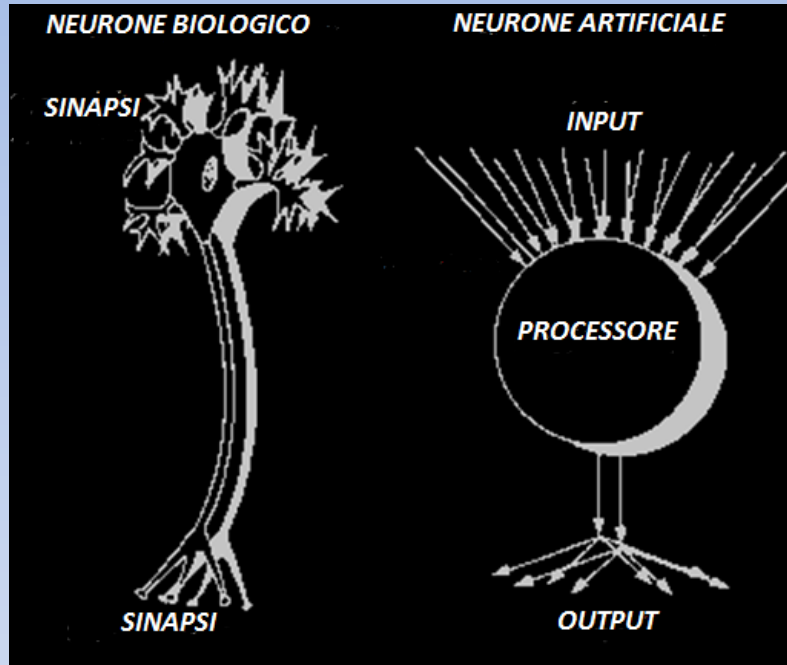
Un neurone può avere **migliaia di punti di ingresso** e può ricevere quindi messaggi da un numero straordinario di altre cellule. Ciò rende possibile la trasmissione e l'elaborazione di informazioni complesse. I neuroni hanno la possibilità di integrare segnali eccitatori e inibitori che ricevono in ogni momento. Quando la somma algebrica dei potenziali locali supera il valore di soglia, il neurone genera un impulso nervoso (**potenziale d'azione**). Nella maggior parte dei neuroni la sommazione avviene nel **cono di emergenza**, la regione del corpo che si trova alla base dell'assone.

Se un neurone può ricevere un numero altissimo di input sinaptici, tuttavia fornisce un solo e univoco output di uscita che viene distribuito alla rete a cui è connesso.

IL NEURONE ARTIFICIALE

NEURONI NATURALI E ARTIFICIALI

STORIA



Il prototipo esistente in natura di una «macchina che impara» è il cervello come sistema costituito da minuscole unità, i neuroni, che sono le «celle» o i «nodi» di un vasto sistema di acquisizione, elaborazione e trasmissione di informazioni.

Se volessimo individuare una data di nascita dell'attuale analisi computazionale del cervello potremmo citare il libro di **Kenneth Craik** *The nature of explanation* (1943). Il sistema nervoso viene paragonato ad una macchina computatrice in grado di modellare e di mettere in parallelo gli eventi esterni.

La neurofisiologia è un riflesso che serve ad **aggiustare il corpo nelle sue risposte ad un mondo esterno**. La conoscenza si sviluppa all'interno di un rapporto circolare, dinamico e interattivo tra l'organismo e l'ambiente in un quadro di tipo adattativo. All'interno di questo circolo viaggia una massa di informazioni /segnali che passano dall'ambiente all'organismo e che vengono selezionate e distribuite e processate su una molteplicità di canali in un contesto analitico e differenziato. L'intero flusso infine confluisce nel cervello dove le informazioni si traducono in un ricchissimo alfabeto neurochimico e vengono a loro volta ricomposte e assemblate in modo da consentire all'individuo di procurarsi un ritratto della realtà entro il quale ottimizzare il suo rapporto con l'ambiente.

RETI NEURALI

La **rete neurale** «naturale» è un sistema «elastico» che è in grado di modificare la sua struttura al fine di ottimizzare le sue risposte o prestazioni ambientali. Le sinapsi modificano le loro proprietà al fine di fornire le **risposte** più adeguate agli stimoli che provengono dal mondo esterno (o interno).

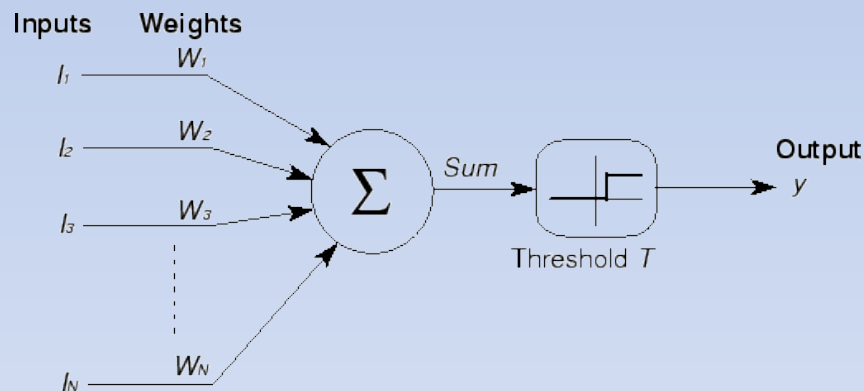
Le **reti neurali artificiali** cercano di **emulare** questo comportamento elaborando **modelli computazionali**, più o meno complessi, che consentano di descrivere e organizzare un sistema in grado di adottare strategie **autonome** di assestamento e controllo dei propri ingressi al fine di uniformare l'uscita ad un risultato desiderato.

Il concetto di «peso». Ciascun «neurone» o nodo della rete è caratterizzato da un numero elevato di canali di ingresso e di un unico canale di uscita. Il segnale di uscita è o un segnale qualificato generato dalla «somma» dei dati che confluiscono nel nodo. *Ad ogni ingresso (connessione) viene attribuito un «peso» che consiste nel contributo individuale, positivo o negativo, reso dal singolo ingresso nel determinare il risultato finale.* La somma algebrica di tutti i pesi non viene passata come tale al canale di uscita, ma viene a sua volta elaborata da una particolare funzione algebrica detta *funzione di trasferimento*.

Implementazione del concetto classico di «intelligenza artificiale». Anche intuitivamente si comprende come tale sistema ispirato ai modelli biologici **vada oltre** il modello computazionale classico dei calcolatori seriali che pur sono in grado di emulare nostri processi cognitivi superiori (si pensi al computer che gioca a scacchi), ma che non saprebbero ad esempio processare immagini visive.

IL NEURONE DI McCULLOCH-PITTS

Nel 1943 Warren S. McCulloch, neurologo, and Walter Pitts, matematico, pubblicarono *A logical calculus of the ideas immanent in nervous activity*. In questo studio McCulloch e Pitts cercarono di dimostrare come il cervello potesse produrre modelli altamente complessi usando una molteplicità di cellule di base (neuroni) interconnesse insieme. McCulloch and Pitts produssero un modello semplificato di neurone.



Questo modello-base di neurone ha costituito un importante contributo per lo sviluppo delle reti neurali– con caratteristiche simili al funzionamento della rete nervosa. I neuroni di MacCulloch-Pitts avevano però dei limiti. Caratteristiche ulteriori furono in seguito aggiunte per consentire al neurone di “imparare”. Il successivo passo decisivo per lo sviluppo delle reti neurali fu il concetto di perceptrone introdotto da Frank Rosenblatt nel 1958.

Ogni neurone ha molte connessioni in ingresso e una sola in uscita che a sua volta costituisce una porta di ingresso di un altro neurone. L’uscita assume un valore 1 se la somma dei pesi associati alle singole connessioni supera un determinato **valore di soglia**, altrimenti il valore è 0.

L'algoritmo di apprendimento specifica in che modo debbano essere modificati i pesi per far sì che la rete ... «apprenda». Cosa significa «apprendere» per una macchina? Però immaginiamo di trovarci di fronte ad una superficie maculata di punti e figure casuali. Qualcuno vi vede riflessa l'immagine di Padre Pio. Qualcun altro vi intravede una lettera dell'alfabeto. Come possiamo insegnare ad una macchina a riconoscervi dei volti, una lettera dell'alfabeto, il profilo di una montagna, classificare una immagine ... ?

Il sistema deve essere «educato» (addestrato) a riconoscere configurazioni e a reagire mettendo in atto **algoritmi di aggiustamento** a fronte di una configurazione equivoca. Il flusso di ingressi all'inizio assume un assetto casuale. Attraverso un *training* di addestramento il sistema impara a modificare i pesi – ossia a classificare, ordinare, modificare ecc. – i dati di ingresso al fine di ridurre progressivamente la differenza tra il risultato attuale e quello desiderato.

I pattern utilizzati nella fase di apprendimento sono chiamati pattern di addestramento. La forma della risposta della rete, la configurazione di uscita, può essere controllata *a posteriori* o *a priori*.

Apprendimento non supervisionato. Significa che l'algoritmo prevede una valutazione delle varie configurazioni dei dati di ingresso. Sulla base delle proprietà dei campioni di ingresso, l'algoritmo modifica i pesi in modo da **implementare** questo o quel modello di uscita.

Apprendimento supervisionato. Alla rete è necessario fornire degli «esempi» e la modifica dei pesi avviene in base al **confronto** fra uscita effettiva della rete (sbagliata, all'inizio) e uscita desiderata. ¹⁸

CONTROLLI *FEED-BACK* E *FEED-FORWARD*

Sono due concetti importanti nel panorama che andiamo esplorando.

Feed-back. E' un tipo di valutazione che viene fatta a posteriori dopo avere ottenuto il risultato. Ad esempio un termostato compara i dati di *input* (la temperatura ambientale) con un *target* (esempio: impedire che la temperatura scenda al di sotto di 13 gradi). Quando la temperatura raggiungerà i 12 gradi, il termostato accenderà la caldaia. Però in tal caso **la correzione avviene quando l'errore è già avvenuto!**

Feed-forward. Ovviamente la soluzione ottimale consisterebbe nell'intervenire **prima** che la stanza finisca al freddo, attraverso una sapiente gestione e interpretazione degli input sensori che consenta alla macchina di **anticipare** l'errore, l'evento negativo. Questa capacità previsionale viene acquisita dal sistema elaborando esperienze precedenti attraverso un *training* di apprendimento associativo in cui i segnali di errore vengono utilizzati per aggiustare i parametri decisionali e correggere gli output devianti. Immaginando un termostato che funzioni con questo tipo di controllo, esso sarebbe in grado di valutare e giostrare con i tempi di accensione della caldaia, sulla base di statistiche precedenti relativi alle variazioni della temperatura in modo da **prevenire** la discesa della colonnina del mercurio al di sotto dei 13 gradi.

Una rete neurale è una tipica rete di controllo *feed-forward*. La sua versione più semplice è il perceptrone .

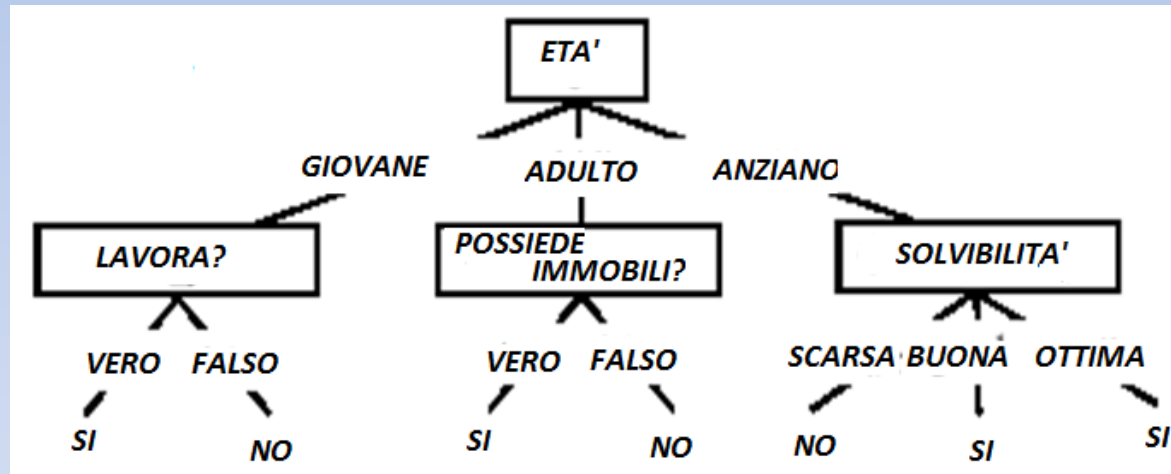
L'ALBERO DI DECISIONE

L'**albero di decisione** costituisce la base più semplice di un algoritmo di apprendimento. I dati vengono classificati e ordinati secondo le loro caratteristiche di valore che diventano sempre più specifiche fino a portare al «risponso finale», ossia al SI o al NO.

Immaginiamo un robot che di professione faccia il banchiere e debba quindi concedere prestiti, imparando a valutare se il cliente sia affidabile o non affidabile e in grado di restituire il prestito (obiettivo desiderato). Il robot deve «fare esperienza» valutando alcuni parametri.

Nel nostro esempio i dati di input sono relativamente pochi e basati sulla età della persona.

Dobbiamo pensare a scenari molto più vasti e complessi in grado di estrapolare informazioni e perciò di «apprendere» da una massa enorme di dati (**big data**).



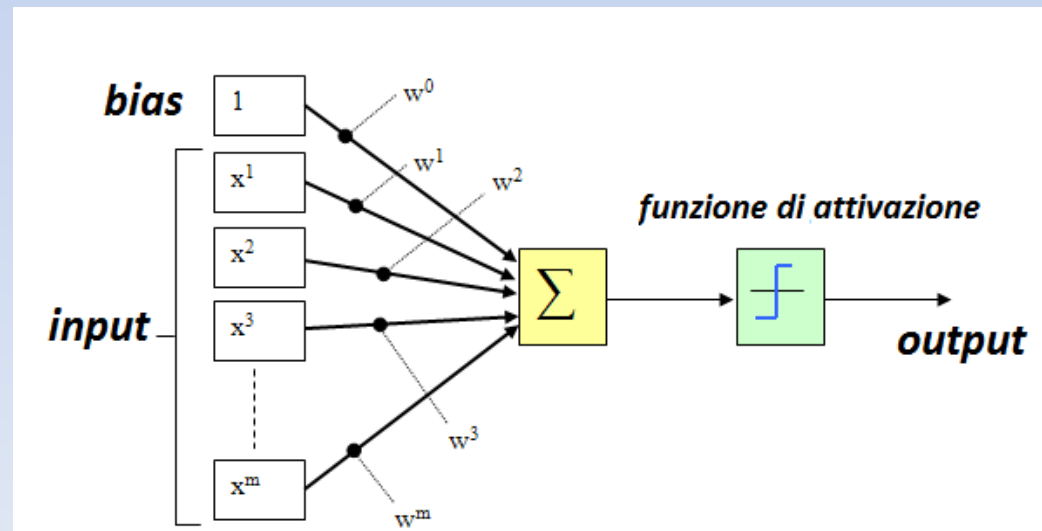
Data mining. Complesso di tecniche computazionali che consentono di estrapolare modelli (o categorie) da un enorme *set* di dati (**big data**). Ciò consente di estrarre da una massa di dati informazioni inizialmente **non conosciute**, e rese disponibili per ulteriori utilizzi. Queste procedure rappresentano una fase importante nella architettura ideale di una *learning machine*, che in questo modo riesce a classificare i dati a seconda di parametri di affinità, omogeneità, ricorrenza, discostamento ecc.

IL PERCETTRONE di ROSENTHAL

In filosofia e in psicologia “*per-cepire*” significa ... *prendere due volte* (e forse anche di più). La percezione si distingue dalla mera sensazione perché a questa si sommano anche componenti e controlli di origine soggettiva – *apriori* o derivati da esperienze precedenti . Il perceptrone riproduce questa stessa modalità **apprendimento supervisionato**.

Essenzialmente il perceptrone è un neurone di MacCulloch-Pitts integrato da una **funzione** che gli consente di **classificare** il flusso dei dati in ingresso associandoli ad un determinato tipo di uscita. Qualora la configurazione di ingresso non produca l’uscita desiderata, il sistema è in grado di “tornare indietro” (**retropropagazione**) a modificare il flusso dei dati in ingresso fino ad ottenere l’uscita desiderata (regola di *minimizzazione dell’errore*).

Nello schema più semplice, per alterare il peso delle connessioni in ingresso dall’esterno, è possibile agire inserendo un peso sinaptico fittizio (interno) detto **bias**

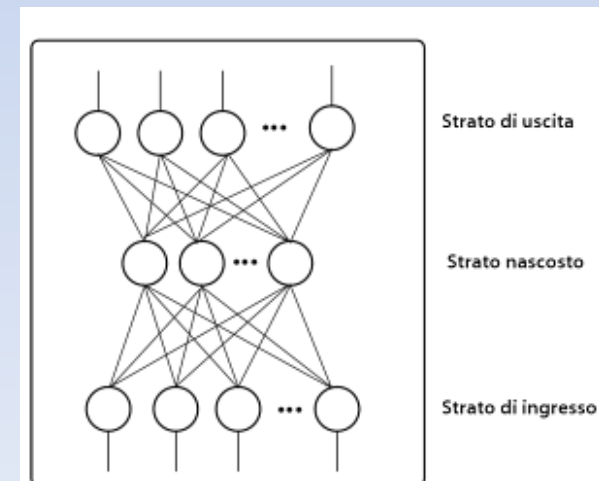


RETI MULTISTRATO

Limiti del perceptrone. Il perceptrone ha rappresentato una tappa fondamentale per lo sviluppo della cibernetica. Tuttavia esso non è in grado di operare, nella sua fase di addestramento, delle classificazioni particolarmente complesse, gestendo separatamente alcune caratteristiche dell'oggetto per poi assemblare il risultato.

Una **rete multistrato** supera questa limitazione in quanto può creare **rappresentazioni interne** (nascoste) e *imparare differenti caratteristiche* in ciascuno strato. Ad esempio, dovendo leggere una immagine, il primo strato può essere responsabile dell'apprendimento dell'orientamento delle linee usando gli input dai singoli pixel nella immagine. Il secondo strato può combinare le caratteristiche apprese nel primo strato e imparare a identificare semplici figure come ad esempio un circolo o un triangolo ecc. *Ogni strato superiore apprende caratteristiche sempre più astratte che possono infine essere poi assemblate per rappresentare l'intera immagine.*

Mentre il perceptrone «impara» con poche prove di apprendimento, queste reti, visibilmente più complesse, richiedono centinaia di prove per imparare la giusta corrispondenza tra ingresso e uscita. Ciò comporta un problema in robotica (se un robot deve evitare un ostacolo, dovrebbe ripetere la tipologia di errore centinaia di volte per evitarlo).



RETI HEBBIANE

*(Neurons that fire together
wire together)*

Donald Hebb rappresenta una figura di primo piano e un grande anticipatore nella storia della psicobiologia cognitiva (*The organization of behavior. A neuropsychological theory, 1949*). Quando oggi parliamo di “connettomica” o della possibilità di ricostruire per intero in forma di mappe cerebrali la fenomenologia dei nostri comportamenti e processi mentali, dobbiamo ripensare a Hebb.

Legge di Hebb (apprendimento associativo): *La simultanea attivazione di cellule neuronali provoca un significativo rafforzamento del loro legame sinaptico e fornisce una base biologica per processi di apprendimento liberi da errori e utili ai fini dell'addestramento e della riattivazione delle memorie.*

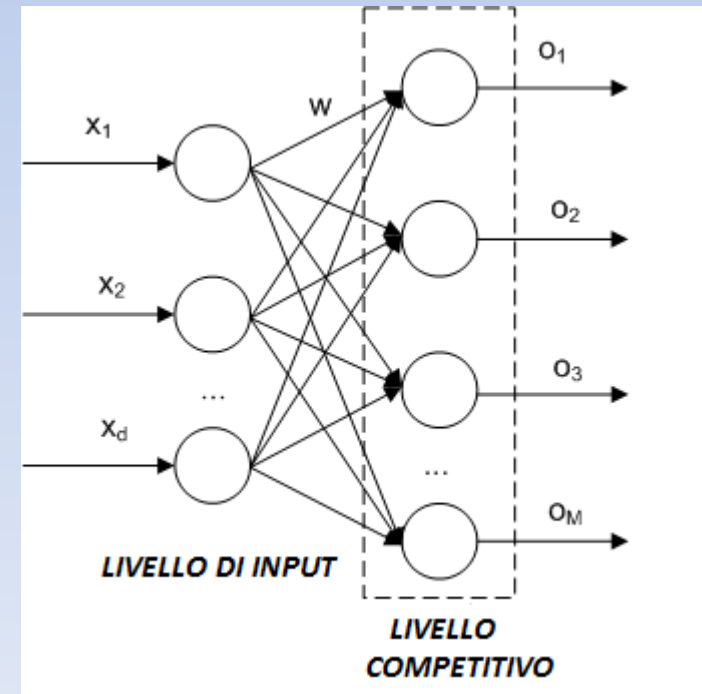
Questo principio, trasferito su un più vasto scenario reticolare, fornisce un metodo per la **gestione dei pesi** all'interno di modelli neuronali. *Il peso tra due neuroni aumenta se i due neuroni si attivano simultaneamente e si riduce se essi si attivano separatamente.* Nodi di rete che tendono simultaneamente ad essere entrambi positivi (o entrambi negativi) hanno peso superiore rispetto a quelli che si comportano in modo opposto.

RETI COMPETITIVE -1

(WINNER MAKE ALL)

È una forma di apprendimento (**non supervisionato**) nelle reti artificiali neurali dove i nodi **competono** per acquisire il diritto di rispondere ad un sottogruppo di *input*. Sotto questo aspetto la rete riesce ad emulare il concetto di scelta o di **decisione**. I nodi di *output* si inibiscono reciprocamente e dopo un lasso di tempo solo un nodo nella progettazione di *output* rimane attivo. Praticamente ogni neurone elabora in modo diverso gli *input* in ingresso, ma infine solo una di queste elaborazioni diventa vincente. Ciò consente di estrarre di volta in volta dalla rete le soluzioni più funzionali.

La rete assume una configurazione articolata che emula il modello di comportamento delle popolazioni neurali biologiche dove il singolo neurone che si è specializzato a processare un particolare stimolo, inibisce questa facoltà agli altri (inibizione laterale).



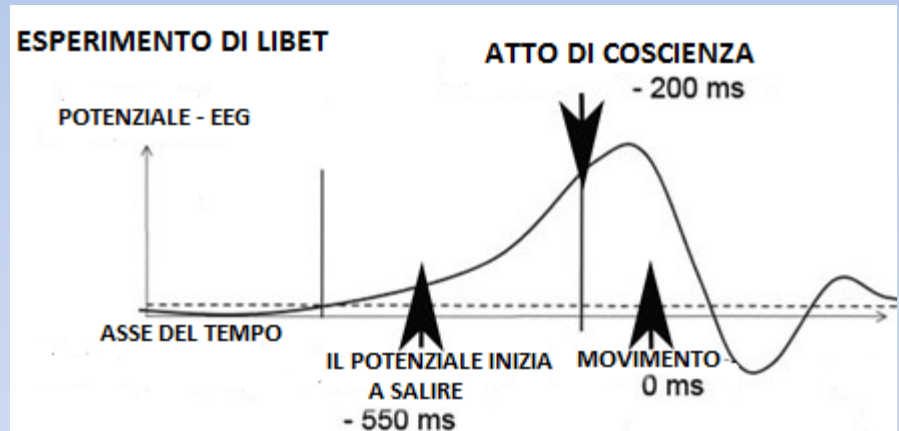
RETI COMPETITIVE - 2

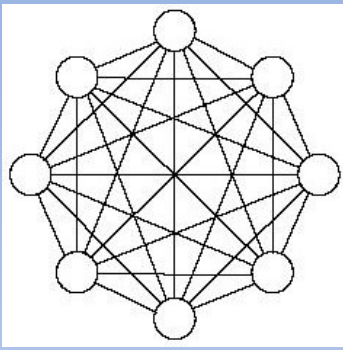
Il problema del *free will*. L'architettura delle reti competitive sembra voler emulare il meccanismo decisionale umano. Le nostre scelte scaturiscono forse da una «tempesta neuronale» (e perciò inconscia) dove infine uno stimolo vince su tutti gli altri? E quale ruolo gioca allora la coscienza?

L'esperimento di Libet. Per certi aspetti, l'esperimento di Libet, ripetuto in seguito da altri, sembra confermare questo scenario in cui le leve che inducono le nostre scelte siano soprattutto inconscie (e cioè «neurologiche»). L'esperimento tendeva a misurare i tempi intercorrenti tra il momento della decisione (ad esempio la decisione di muovere un dito) e la sua esecuzione da parte del cervello.

Sorprendentemente i tempi risultavano invertiti: il cervello iniziava l'azione 350 ms **prima** che il soggetto decidesse di attuarla! Apparentemente era la «rete neuronale» ad aver preso la decisione.

Il ruolo della coscienza. L'esperimento ha suscitato ampie discussioni che ancora proseguono. Siamo «liberi» oppure automi? Dai suddetti esperimenti risulta tuttavia che questo ruolo di subordine della coscienza non si verifica nel caso di una **scelta negativa**, ossia nell'atto di inibizione di una azione. La nostra libertà si manifesta nel dire «no»





RETI DI HOPFIELD

RETI NEURALI

Nel modello McCulloch-Pitts ogni neurone processa i suoi input per determinare i propri output. La rete di Hopfield viceversa è una **struttura associativa** (simile alla interconnessione dei neuroni del cervello umano) per cui dobbiamo parlare di **input di rete** ed **output della rete** e di «stati» o settaggi di rete. Esiste solo un ingresso di rete ed una uscita di rete. Ogni singolo «neurone» esprime un «pezzetto» di possibili configurazioni globali.

Memoria di rete. La natura di queste rete è che essa lavora con modelli di configurazione memorizzati che vengono richiamati in presenza di modelli di ingresso incompleti o «rumorosi». In pratica viene ricostruito il pattern originale da un modello difettoso (*error-correcting*). Dopo l'addestramento su un insieme di esempi, un nuovo stimolo porterà la rete a stabilizzarsi su una configurazione di attivazioni corrispondente all'esempio nell'insieme di addestramento che assomiglia di più al nuovo stimolo.

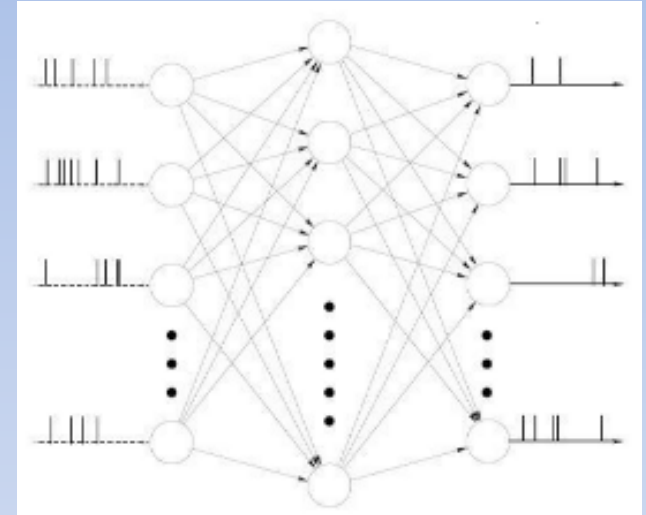
Il limite del modello è dato dal fatto che la rete non è in grado di memorizzare informazioni se il numero dei dati supera il 14% dei neuroni della rete.

RETI NEURALI AD IMPULSI ED IL SOFTWARE *SPIKESTREAM*

RETI NEURALI
ROBOTICA

Le reti neurali ad impulsi (*Spiking neural networks*)

incrementano il livello di realismo nella simulazione neuronale. Rappresentano la **terza generazione** delle reti neurali e si basa su un modello di comunicazione tra neuroni che sfrutta un linguaggio costituito da sequenze di impulsi elettrici – come una specie di sofisticato di linguaggio Morse. Ogni neurone «emette» in fase di output un *treno di impulsi*. In questo modo un neurone, potendo disporre di un linguaggio di comunicazione che consente molte varianti, può svolgere la funzione di centinaia di neuroni «nascosti» in una rete convenzionale multistrato.



Software SPIKESTREAM. Simulatore di reti neurali ad impulsi parte del progetto CRONOS-SIMNOS (*Università di Bristol e Essen*) che consente di editare (creare o cancellare) e gestire un sistema fino a 100.000 neuroni. Il software è scritto in C++. Dotato di interfaccia grafica 3D, possono essere inseriti neuroni e settate sinapsi, collegare apparati esterni e avviate le simulazioni. Le informazioni introdotte vengono memorizzate in un database.

COSCIENZA ARTIFICIALE

NEUROETOLOGIA

LA RANA COMPUTATRIX

Il concetto di «mappa cognitiva». Concetto introdotto da Tolman (1948) riferito in origine allo studio delle capacità animali di pianificare l'esplorazione del territorio. In generale le mappe cognitive sono delle *scorciatoie* che consentono di pianificare dei comportamenti minimizzando gli errori.

Il concetto di schema. Incontriamo termini quali «schema sensoriale», «schema motorio». Che cosa è più precisamente uno «schema»? Unità di assemblaggio di funzioni e circuiti neuronali utile che concorrono a formare un sistema complesso sia sul piano strutturale che funzionale. Uno schema può comprendere molti circuiti neuronali.

Rana computatrix. E' un automa (che imita il comportamento della rana che cattura un insetto) costruito su **4 schemi**: 2 Schemi percettivi per riconoscere oggetti o situazioni e 2 schemi motori per controllare strutture di azione e comportamento. E' parte della più vasta famiglia di modelli computazionali per lo studio della coordinazione visuomotoria. In seguito sono stati proposti altri modelli più complessi, come ad esempio il gesto dell'«afferrare» con la mano. Se tra il dire e il fare «c'è di mezzo al mare» - come da proverbio – viceversa tra il fare e il dire il mare non c'è, per cui i modelli computazionali studiati sul gesto dell'«afferrare» vengono presi a base per lo studio di modelli ancora più complessi che riguardano lo sviluppo del linguaggio (M. A. Arbib, *Rana computatrix to human language: towards a computational neuroethology of language evolution*, 2003)

COSCIENZA E COSCIENZA ARTIFICIALE

La coscienza. Ancora oggi filosofi e psicologi si misurano con il millenario e irrisolto *hard problem* del rapporto mente/cervello. Esistono molte ipotesi elaborate dagli studiosi negli ultimi decenni. E' però convinzione comune che la capacità del cervello umano di produrre prestazioni a livello superiore sia principalmente dipendente dalla vastità, fitta ed altamente differenziata, della **rete neuronale**, alla sua capacità di gestire su circuiti **in parallelo** *input* interni ed esterni qualitativamente e quantitativamente differenziati, alla velocità ed alla estensione delle connessioni.

Quali sono i criteri per giudicare il «livello di coscienza» di una «macchina»? Una macchina non potrà ancora emulare *in toto* la coscienza umana, però ciò non toglie che si possa valutare il suo «grado» di coscienza, anche sulla base della *quantità e qualità di funzioni* che essa è in grado di emulare.

Domanda: una coscienza *simulata* può essere considerata «reale»? Siamo abituati a considerare «reale» una coscienza che alimenta il comportamento dei sistemi biologici. Eppure se noi apriamo la «scatola» biologica, noi vedremmo agire un complesso concerto di processi neuronali che noi riferiamo a «stati di coscienza». Una stessa capacità «produttiva di coscienza» in senso fenomenologico deve pertanto essere teoricamente riconosciuta anche alla macchina dal momento in cui riproduce gli stessi processi. Ma possiamo considerare questa risposta sufficiente?

La coscienza non è solo in grado di «rispecchiare il mondo», ma di rispecchiare anche se stessa: “... è necessario che abbia nel cervello non solo una rappresentazione del mondo e dei vari oggetti che esso contiene, incluso il mio stesso corpo interno alla rappresentazione ... ma in più la rappresentazione dell'oggetto esterno deve interagire con la rappresentazione di me stesso” (S. Ramachandran)