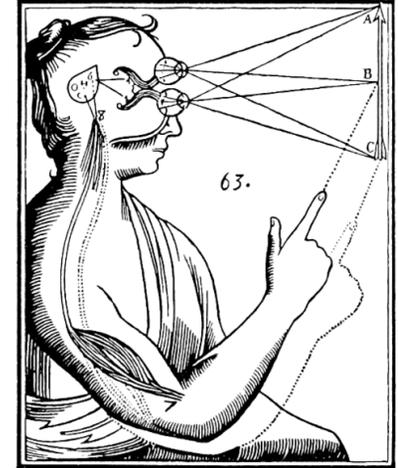


Intelligenza Artificiale: **Origini, sviluppi e punti di vista**

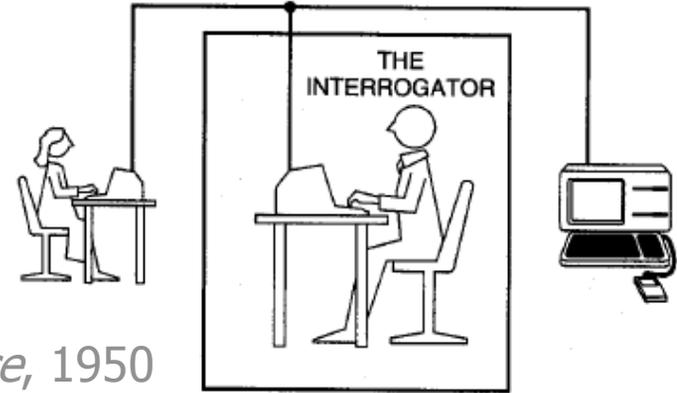
Marco Piastra



- Descartes, R., *Discours de la Methode*, 1637

“[...] avevo descritto l'anima razionale, e fatto vedere che in nessun modo può derivare dalla potenza della materia, come le altre cose di cui avevo parlato, ma che deve essere creata appositamente; e avevo mostrato come non basti che sia posta nel corpo umano, come un pilota nella sua nave, se non forse per muovere le sue membra, ma che bisogna che sia congiunta e unita ad esso ancor più strettamente, perché, oltre a tutto questo, possa provare sentimenti ed appetiti simili ai nostri e costituire in tal modo un vero uomo.”

Can machines think? **(Il test di Turing)**



- Turing, A., *Computing Machinery and Intelligence*, 1950

“The new form of the problem can be described in terms of a game which we call the ‘imitation game.’

It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex.

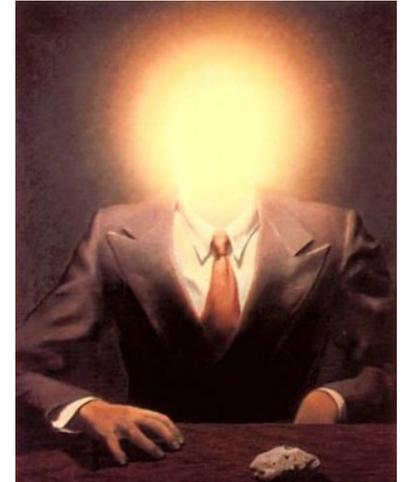
The interrogator stays in a room apart from the other two.

The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either ‘X is A and Y is B’ or ‘X is B and Y is A’

The interrogator is allowed to put questions to A and B. [...]

We now ask the question, ‘What will happen when a machine takes the part of A in this game?’ Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman?

These questions replace our original, ‘Can machines think?’ ”



- Searle, J. R., *Minds, Brain and Science*, 1986

“Because we do not understand the brain very well we are constantly tempted to use the latest technology as a model for trying to understand it.

In my childhood we were always assured that the brain was a telephone switchboard (*‘What else could it be?’*).

I was amused to see that Sherrington, the great British neuroscientist, thought that the brain worked like a telegraph system. Freud often compared the brain to hydraulic and electro-magnetic systems. Leibniz compared it to a mill, and I am told some of the ancient Greeks thought the brain functions like a catapult.

At present, obviously, the metaphor is the digital computer.”

- Crick, F., *The Astonishing Hypothesis*, 1994
“You, your joys and your sorrows, your memories and your ambitions, your sense of personal identity and free will, are in fact no more than the behavior of a vast assembly of nerve cell and their associated molecules.”

Le origini dell'IA

- John McCarthy et al., 1955

“We propose that a two-month, ten man study of **artificial intelligence** be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire.

The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of **intelligence** can in principle be *so precisely described* that a machine can be made to *simulate* it.

[...] It may be speculated that a large part of human thought consists of manipulating **words** according to **rules** of **reasoning** and **rules** of **conjecture**.

[... *An intelligent machine*] would tend to build up within itself *an abstract model of environment* in which it is placed. If it were given a problem, it could first explore solutions within the internal abstract model of the environment and then attempt external experiments.”

The physical-symbol system hypothesis (PSSH)

- Newell, A., Simon, H., *Computer Science as Empirical Inquiry Symbols and Search*, 1976

“A **physical symbol system** consists of a set of entities, called **symbols**, which are physical patterns that can occur as components of another type of entity called an expression (or symbol structure).

Thus, a **symbol structure** is composed of a number of **instances** (or tokens) of symbols related in some physical way (such as one token being next to another).

At any instant of time the system will contain a collection of these symbol structures.

Besides these structures, the system also contains a collection of **processes** that operate on *expressions* to produce other *expressions*:
processes of creation, modification, reproduction and destruction.”

Struttura simbolica (esempio intuitivo)

- Fatti:
 1. Enrico è il genitore di Paolo
 2. Silvia è la genitrice di Giorgio
 3. Silvia è la genitrice di Renato
 4. Enrico è il genitore di Giorgio
 5. Enrico è un essere umano
 6. Paolo è un essere umano
- Regole:
 7. Tutti gli esseri umani sono bipedi senza piume
 8. Due persone diverse che abbiano gli stessi genitori sono parenti
 9. Chiunque abbia come parente un bipede senza piume è contento

Processi

- Giorgio è contento?

[da 9] chi è parente di Giorgio?

[da 8] chi è genitrice di Giorgio? Silvia
di chi è genitrice Silvia? Renato

Renato e Giorgio sono due persone? sì

[conclusione da 8]: Renato è parente di Giorgio.

[continua con 9] Renato è un bipede senza piume?

[da 7] Renato è un essere umano? non è un fatto noto [fallimento]

[torna a 8] di chi altro è genitrice Silvia? Giorgio

Giorgio e Giorgio sono due persone? no [fallimento]

[torna a 8] di chi altro è genitrice Silvia? nessuno [fallimento]

[da 8] chi altro è genitore di Giorgio? Enrico

di chi è genitore Enrico? Paolo

Paolo e Giorgio sono la stessa cosa? no

[conclusione da 8]: Paolo è parente di Giorgio.

[torna a 9] Paolo è un bipede senza piume?

[da 7] Paolo è un essere umano? sì

[conclusione da 7] Paolo è un bipede senza piume.

[conclusione da 9] Giorgio è contento.

- **Fatti:**

1. Enrico è il genitore di Paolo
2. Silvia è la genitrice di Giorgio
3. Silvia è la genitrice di Renato
4. Enrico è il genitore di Giorgio
5. Enrico è un essere umano
6. Paolo è un essere umano

- **Regole:**

7. Tutti gli esseri umani sono bipedi senza piume
8. Due persone che abbiano gli stessi genitori sono parenti
9. Chiunque abbia come parente un bipede senza piume è contento

Sistema simbolico

- Simboli
 - Hanno una *rappresentazione* (fisica) che ne consente la manipolazione
 - Possono avere *relazioni* con altri simboli (espressioni, regole)
- Elaborazione
 - Il sistema di elaborazione utilizza *simboli e relazioni* ma opera solo sulla dimensione *fisica* (non 'conosce' il *significato** dei simboli)
- Processi
 - L'implementazione (la macchina) è irrilevante
 - Il processo di ragionamento è definito come attività sui simboli (*disembodiment* = il pensiero senza pensatore?)
- Rappresentazione interna:
 - L'insieme dei simboli e delle espressioni in un dato istante

(*qui inteso come relazione tra simboli e mondo esterno)

Caratteristiche fondamentali

- Astrazione

- I simboli e le espressioni possono avere un significato generale (p. es. le regole 7, 8, e 9 dell'esempio precedente)

- Finitismo

- Un sistema simbolico può operare con un numero finito di simboli e regole (eventualmente combinabili in un numero infinito di espressioni)

- Esempio: calcolo simbolico in matematica

$$x^2 + ax + b = 0$$

$$x^2 + 2(a/2)x + a^2/4 - a^2/4 + b = 0$$

$$(x + a/2)^2 - a^2/4 + b = 0$$

$$(x + a/2)^2 = a^2/4 - b$$

$$x = -a/2 \pm (a^2/4 - b)^{1/2}$$

- Una serie di passaggi, di applicazione di regole formali sulle espressioni
- I simboli x , a e b rappresentano un'infinità (non numerabile) di oggetti
- L'equazione, nelle varie forme, esprime un legame di carattere generale

Calcolo simbolico automatico

- **Rigore formale**
 - Simboli, regole di combinazione e processi devono essere definiti in modo preciso e rigoroso (il sistema deve operare in modo autonomo)
- **Semantica**
 - Il rapporto tra simboli e mondo esterno deve essere definito in modo non ambiguo (il sistema non è isolato ma interagisce con l'ambiente)
- **Computabilità**
 - Quali problemi di calcolo simbolico sono risolvibili in modo automatico?
 - Esempio: *equazioni diofantine* (= a valori interi)
non esiste una procedura generale per il calcolo delle soluzioni
- **Efficienza**
 - Quali problemi di calcolo simbolico sono risolvibili in tempo ragionevole e con memoria limitata?
 - Ad esempio in tempo e spazio polinomiali

Un sistema simbolico: Deep Blue

- Breve storia
 - Nel 1945 A. Turing cita il gioco degli scacchi come un esempio di un'attività che le macchine potrebbero un giorno svolgere
 - Nel 1950 A. Turing definisce il primo algoritmo per il gioco degli scacchi
 - Nel 1997 il sistema Deep Blue di IBM batte l'allora campione mondiale Gary Kasparov
- **Deep Blue, 1997** (Campbell, M., Hoane, A. J., Hsu, F., 2001)
 - 30 processori convenzionali (120Mhz) + 480 processori speciali ('chess search engines', ciascuno valuta 2.5M mosse al secondo)
 - Il gioco degli scacchi ammette circa 30 mosse legittime, con 1000 diramazioni in media ad ogni mossa e contromossa
 - Architettura hardware a tre livelli, 30 GB di RAM complessiva
 - Funzioni di valutazione delle mosse in hardware, software in C
 - Utilizzo di un ampio database di partite di grandi maestri (umani)
- Domanda: Deep Blue è intelligente?

Pensiero ed azione

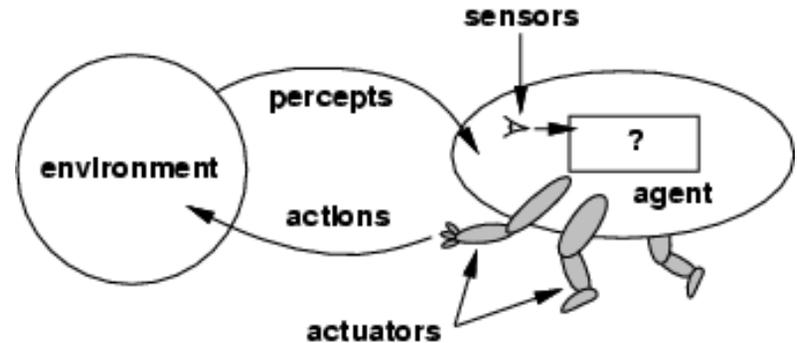
- Otto definizioni (più recenti) di IA

<p>“The exciting new effort to make computers think ... <i>machines with minds</i>, in the full and literal sense” (Haugeland 1985)</p> <p>“The area of computer science that deals with the ways in which computers can be made to perform cognitive functions ascribed to humans” (Noyes 1992)</p>	<p>“The study of mental faculties through the use of computational models” (Charniak and McDermott 1984)</p> <p>“The study of the computations that make it possible to perceive, reason, and act” (Winston 1992)</p>
<p>“A field of study concerned with designing and programming machines to accomplish tasks that people accomplish with their intelligence.” (Schuster 1987)</p> <p>“The study of how to make computers do things at which, at the moment, people are better” (Rich and Knight 1991)</p>	<p>“A field of study that seeks to explain and emulate intelligent behavior in terms of computational processes” (Schalkoff 1990)</p> <p>“The study of intelligent behavior” (Genesereth 1988)</p>

- Classificazione delle definizioni (Russell, S., Norvig, P., 1995)

Sistemi che pensano come gli umani	Sistemi che pensano in modo <i>razionale</i>
Sistemi che agiscono come gli umani	Sistemi che agiscono in modo <i>razionale</i>

Agenti e razionalità



- Agenti
 - Considerare il sistema come un **agente** che interagisce con l'ambiente
 - Elaborazione ciclica (sequenziale o parallela) $f: P^* \rightarrow A$
 - a) Percezione (attraverso i sensori, P^*)
 - b) Valutazione (processi interni $f: P^* \rightarrow A$)
 - c) Azione (attraverso gli attuatori, A)
- Processi interni
 - Quale tipo di processi interni?
 - Come si determina la razionalità dell'agente?
 - Criteri:
 - "Fai la cosa giusta"
(Regole precise di valutazione del comportamento)
 - Calcolo simbolico (forme di logica simbolica)
 - Sviluppare codice 'ad hoc' (i.e. *the hacker's way*)

Razionalità, immaginazione e caso

- (ancora da John McCarthy et al., 1955)

“A fairly attractive and yet clearly incomplete conjecture is that the difference between **creative** thinking and **unimaginative** competent thinking lies in the injection of some **randomness**.

The randomness must be guided by intuition to be efficient. In other words, the educated guess or the hunch include controlled randomness in otherwise orderly thinking.”

Gli elefanti sanno giocare a scacchi?

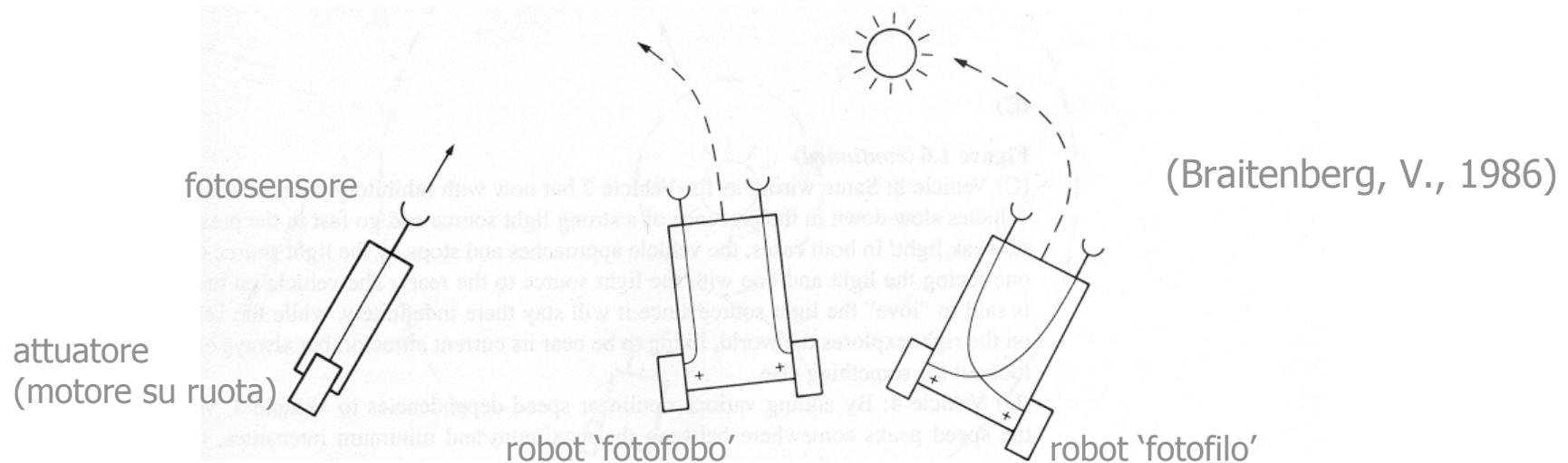
- (Brooks, R., *Elephants Don't Play Chess*, 1990)
- Critica del sistema simbolico (PSSH)
 - Un modello 'monolitico' di rappresentazione interna
 - Gli esseri umani non fanno così – p. es. fenomeno della *change blindness* (O'Reagan, J. K., Rensink, R. A., Clark, J. J., 1999)
 - Un unico sistema (sincrono) di controllo
 - Gli studi sulle lesioni cerebrali suggeriscono il contrario
 - Un sistema di calcolo *general purpose* ed imparziale
 - Lo studio della percezione visiva rivela la 'propensione' a preferire alcune interpretazioni rispetto ad altre
 - Totale separazione di mente e corpo (*disembodiment*)
 - Si disconoscono inoltre le forme di intelligenza non simbolica
 - (Come avrà potuto evolversi, un simile forma di intelligenza?)

Intelligenza come proprietà emergente

- Si manifesta attraverso le interazioni
 - con l'ambiente circostante
 - con altri agenti in una società
 - in base alle conoscenze ed alla storia collettiva
- **Somma di diverse abilità**
 - **Sviluppo**, capacità di acquisire nuove conoscenze
 - **Interazione sociale**, trasferimento ed aiuto reciproco
 - **Corporeità** (*embodiment*) e **contesto fisico** (*physical coupling*)
 - Uso del corpo e dell'ambiente circostante come strumento
 - **Integrazione**, tra abilità e comportamenti diversi

 - 'Patchwork of behaviors' (Brooks, R., 1990)
 - Stratificazione gerarchica (*subsumption*) di comportamenti che legano percezioni ad azioni

Comportamento ed ambiente

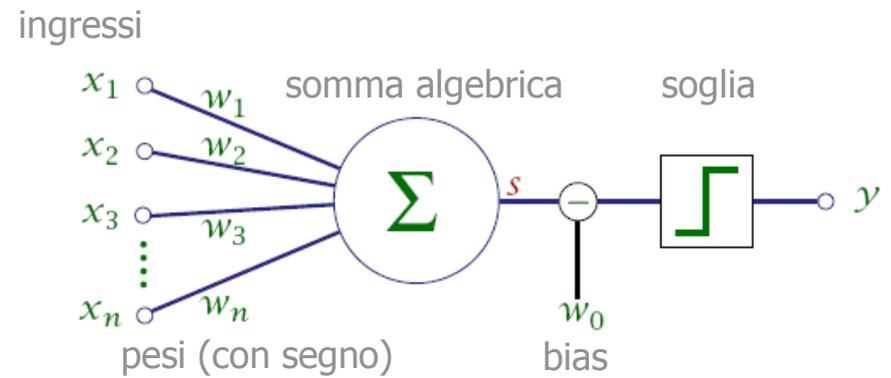
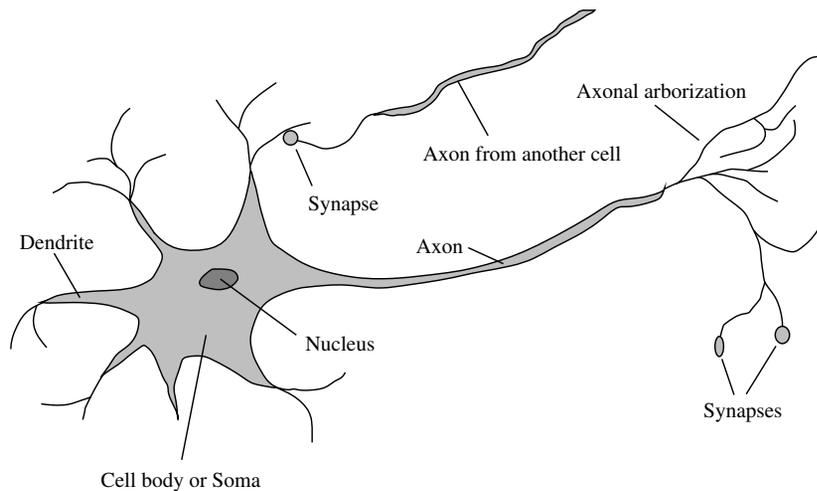


- Connessione diretta

- I robot (p.es. fotofili) che si muovono in un ambiente accidentato mostrano un comportamento (percorso) molto complesso e difficile da descrivere (vedi anche esempio della formica sulla spiaggia - Simon, H., 1981)
- Tuttavia, non possiedono alcuna **rappresentazione interna** dell'ambiente e non hanno capacità di elaborazione simbolica

Connessionismo

- “Occorre considerare anche il *cervello*, non solo la mente”
 - un’ipotesi opposta al *disembodiment* tipico del ragionamento simbolico
- Un modello derivato dallo studio della biologia
 - Nel 1943, McCulloch e Pitts descrivono un modello computazionale formato da una rete interconnessa di neuroni artificiali definiti da un modello a soglia binaria.
 - Dimostrano che il comportamento della rete può essere descritto in termini di calcolo simbolico (logica del primo ordine)



Reti neurali artificiali

- Processi sui segnali

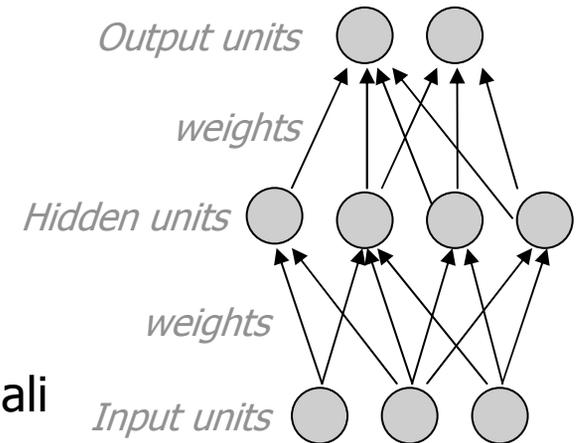
- L'elaborazione si basa sulla propagazione dei segnali attraverso la rete, a partire dalle unità di input
- Qualsiasi automa a stati finiti può essere tradotto in una rete neurale (con *feedback*, vedi oltre) (Minsky, M., 1967)

- Divergenza dal sistema simbolico

- Il modello è fondato sulla fisicità dei **segnali** e non dei simboli
- Il simbolo esiste solo nell'ambito del *comportamento osservabile* (p. es. una configurazione di segnali emessi dalle unità di output)
- Non esiste una rappresentazione interna immediatamente riconoscibile
- Elaborazione intrinsecamente parallela

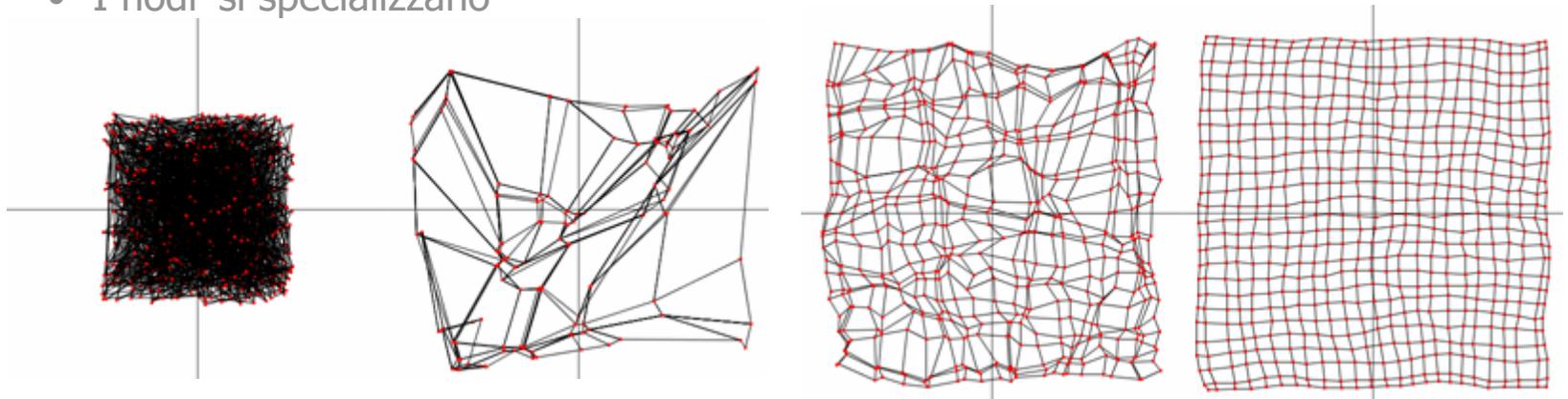
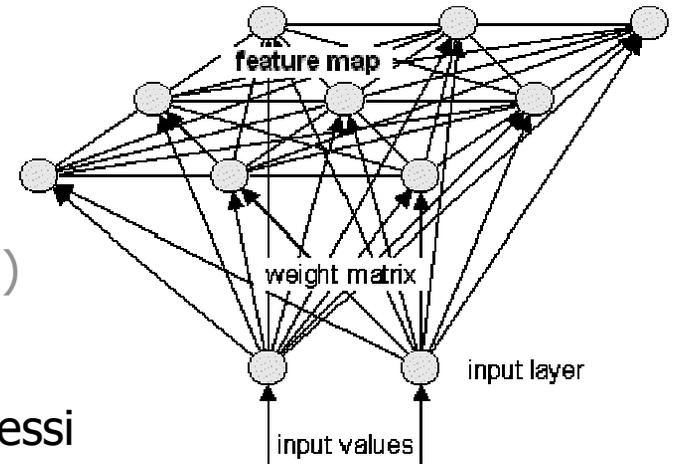
- Apprendimento supervisionato

- Intorno al 1986 viene scoperto un metodo per l'apprendimento supervisionato (*backpropagation*)
 - Le reti apprendono le configurazioni dei pesi a partire da associazioni input/output note a priori



Organizzazione spontanea

- (Kohonen, T., *Self-Organizing Maps*, 1995)
- **Struttura (tipica) a due livelli**
 - Livello di input, con le unità connesse agli ingressi
 - Livello di mappa, con unità organizzate secondo una topologia prestabilita
 - Gli input modificano i pesi secondo un criterio di prossimità
- **Apprendimento non supervisionato**
 - Le SOM si adattano progressivamente alla 'forma' del segnale in input
 - I nodi 'si specializzano'



Feedback e reti ricorsive

- Reti neurali *feedforward* e *feedback*

- Le reti neurali *feedforward* realizzano **funzioni**

- *Universal Approximators* (Hornik et al., 1989)

Comunque traducibili in una rete con un solo livello di *hidden units*

- Per simulare un automa a stati finiti, è necessario il *feedback*

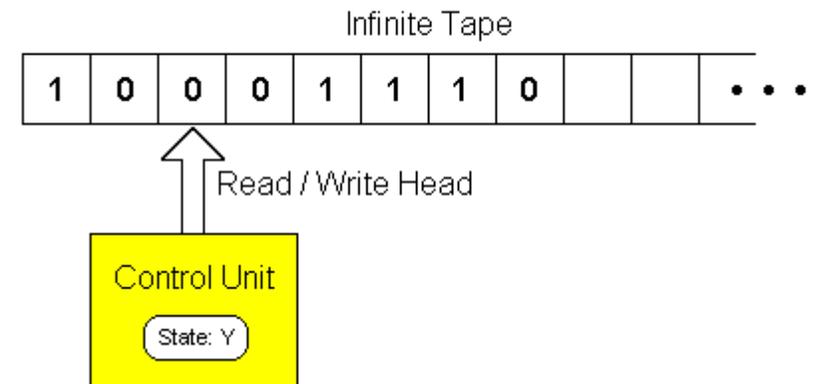
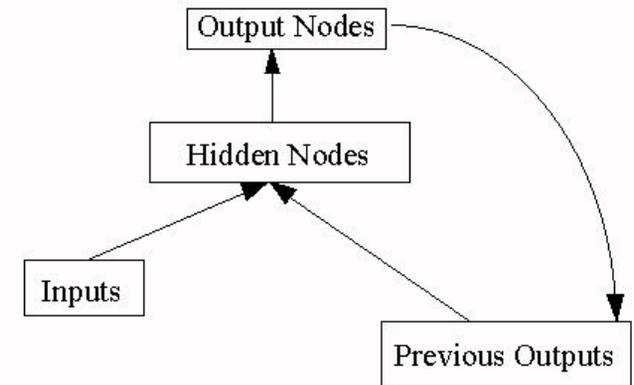
- Per definizione, una funzione descrive un sistema *stateless*

- Reti neurali ricorsive

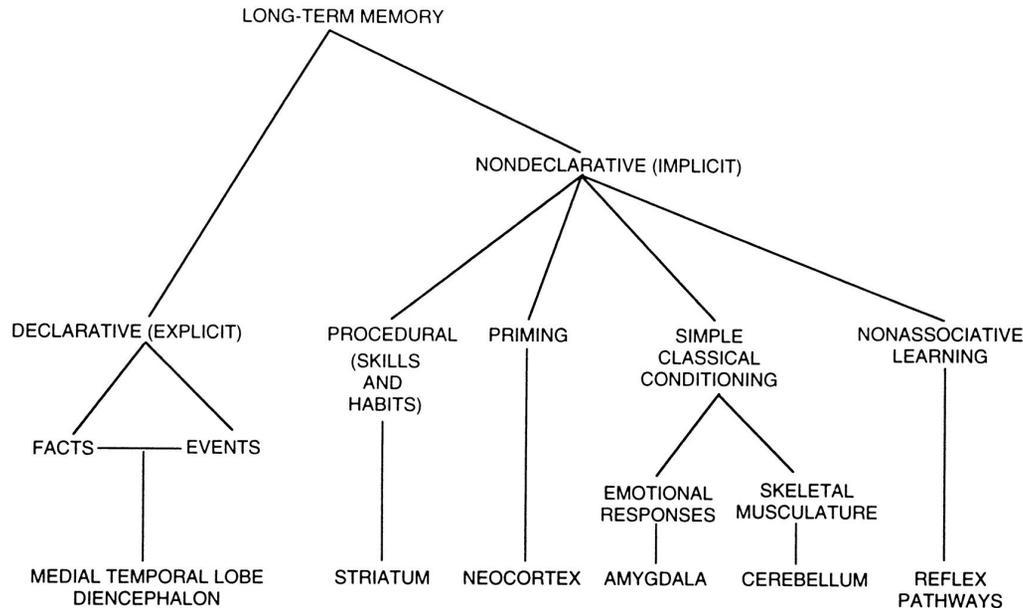
- Il feedback include unità di memoria nel tempo
(il requisito minimo è avere delle *delay unit*)
- Sono traducibili in *Turing Machines*
(Hyötyniemi, H., 1996)

- Complessità

- Non si conosce un metodo generale di apprendimento per le reti neurali ricorsive



Memoria implicita ed esplicita



(Squire, L. R., Zola, S. M., 1996)

- La memoria di lungo termine è costituita da moduli distinti
 - Le lesioni provocano il degrado di capacità diverse a seconda della zona colpita
- Distinzione fondamentale tra memoria dichiarativa e non dichiarativa
 - La memoria non dichiarativa è il frutto dell'apprendimento di azioni
 - La memoria dichiarativa è il consapevole ricordo di fatti ed eventi del passato

Comportamento e memoria esplicita

- Comportamento inconsapevole e rappresentazione cognitiva
 - Le interazioni possono influenzare direttamente il livello neuro-motorio
- Esempio: *visual search*
 - Gli esseri umani effettuano piccoli movimenti a scatti con gli occhi, ad intervalli brevissimi e variabili (tipicamente ~50 msec)
 - Non ne siamo consapevoli: la cosa fu studiata solo alla fine degli anni '60 (Yarbus, A., *Eye Movements and Vision*, 1967)

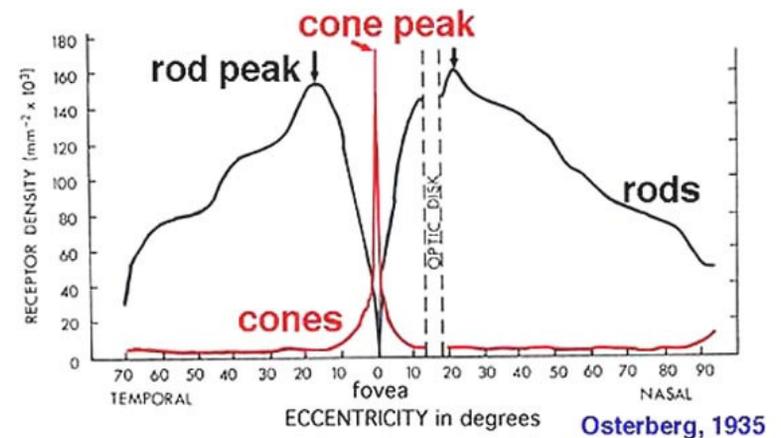
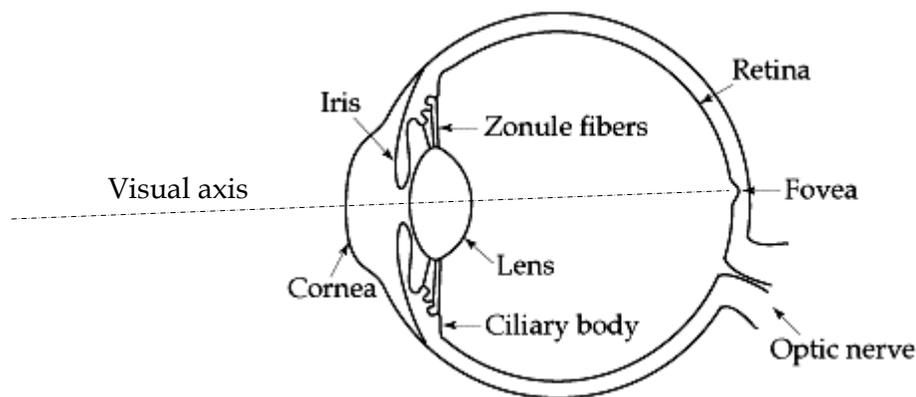
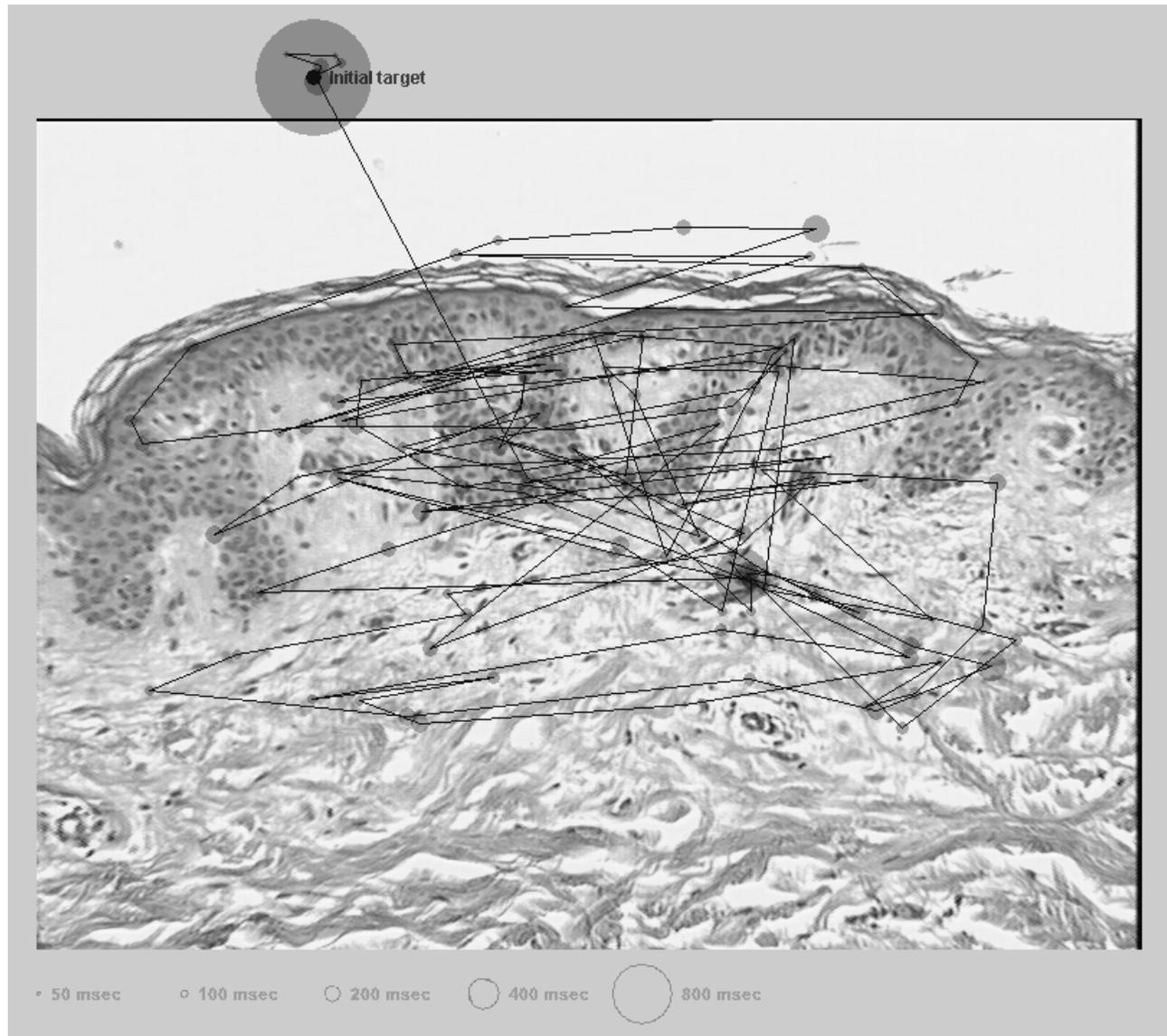


Immagine al microscopio di una sezione di cute opportunamente trattata

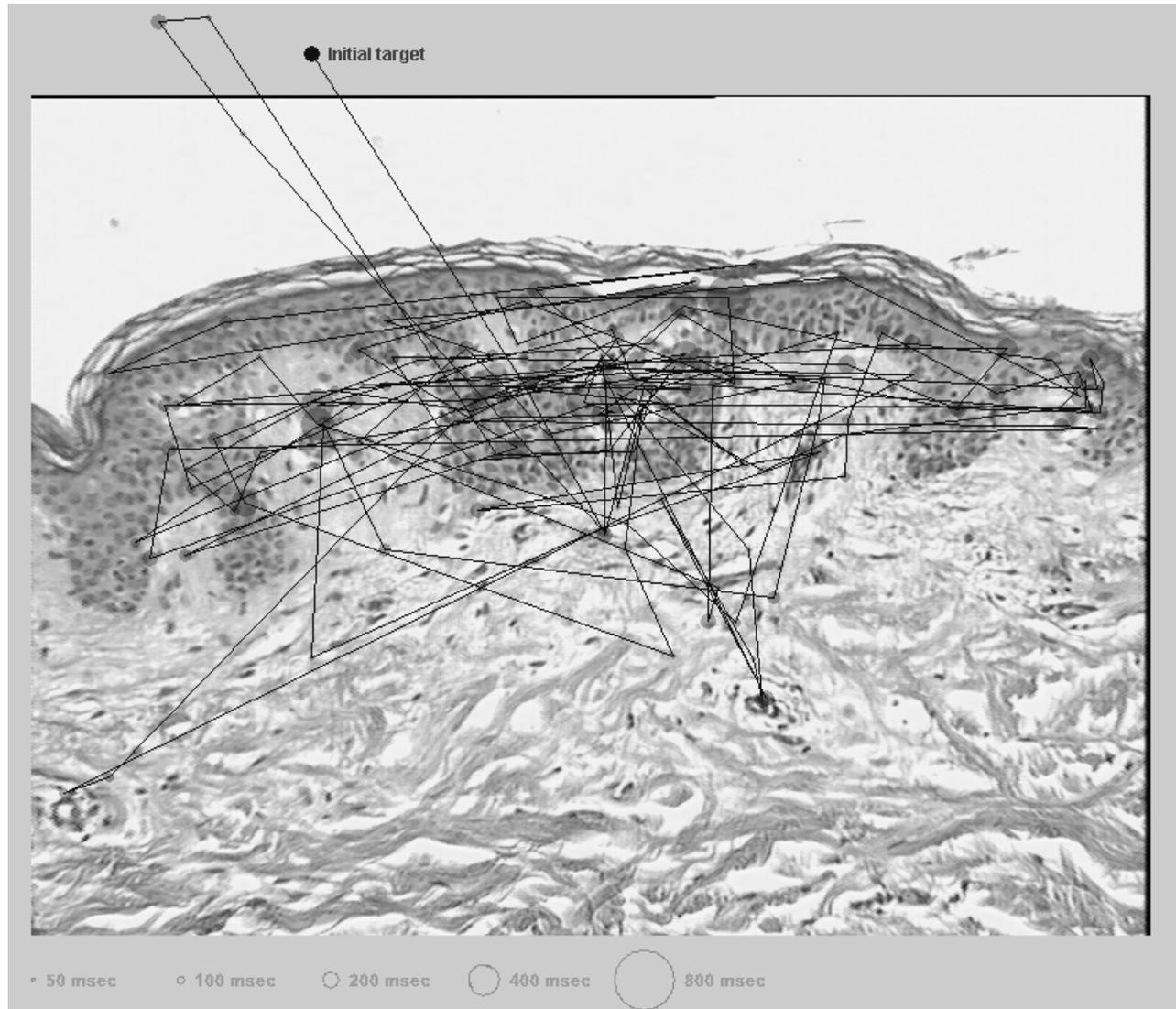
Lo scopo dell'osservazione è una *diagnosi*

I soggetti vedono le immagini per la prima volta

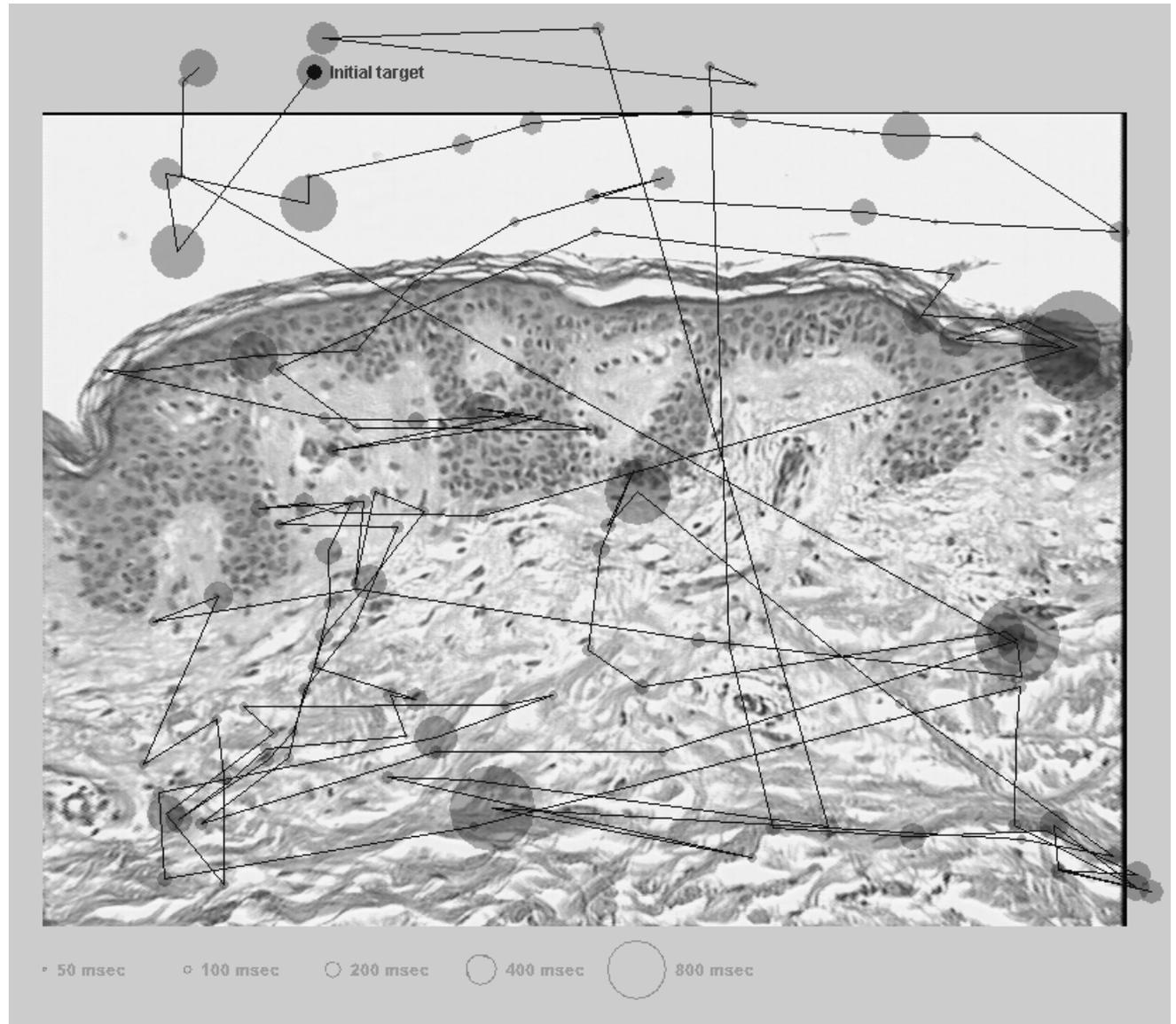
Lo *scanpath* qui tracciato è di un giovane medico (durata: 45 sec)



Lo *scanpath*
di un medico
molto esperto



Lo *scanpath* prodotto da un ingegnere (non conosce l'argomento)



Sistematicità

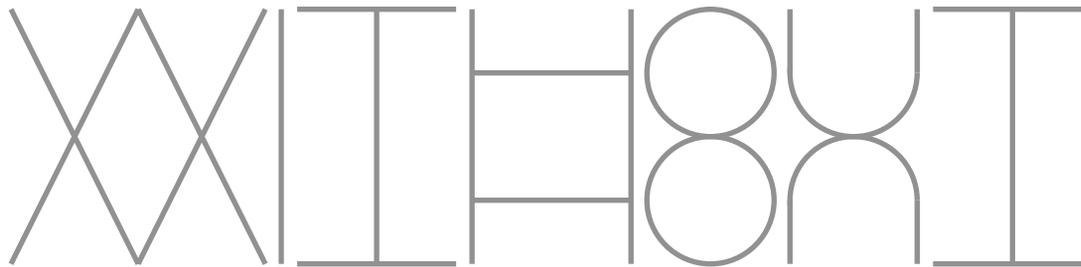
- Senza ricorrere alle capacità di astrazione della descrizione simbolica è difficile spiegare la *sistematicità* di molti fenomeni linguistici (“non è un semplice pattern-matching”)
 - Una persona che comprende l’italiano non può comprendere:
 - “Silvia ama Enrico”senza al tempo stesso comprendere:
 - “Enrico ama Silvia”così come qualsiasi frase del tipo:
 - “X ama Y”dove X ed Y possono essere nomi o descrizioni definite qualsiasi:
 - “L’amica di Renato ama il gatto di Paolo”

(adattato da Fodor e Phylyshyn, 1988)

*Un ulteriore esperimento >

Rappresentazioni diverse

- Fenomeni di sistematicità possono ricorrere anche tra forme di rappresentazione diverse



(adattato da Kanisza, G., 1985)

- Una volta capito il 'trucco', qualsiasi forma analoga diventa immediatamente riconoscibile (astrazione)
- E' molto difficile tornare ad uno stato di 'ingenuità', a prima di aver capito
- Il 'trucco' è trasmissibile (spiegabile) anche a parole, senza immagini

Logica e calcolo simbolico

- Centralità
 - Nella concezione attuale dell'IA, la logica simbolica non ha più un ruolo centrale (come accadeva negli anni '80)
 - In particolare per i problemi di percezione, appare evidente che il calcolo simbolico, da solo, non è in grado di fornire le risposte necessarie
- Rilevanza
 - Nessun sistema alternativo tuttavia presenta le stesse potenzialità descrittive e di analisi per:
 - Processi di ragionamento deduttivo, con astrazione ed istanziamento
 - Formulazione e spiegazione di ipotesi razionali
 - Stati di conoscenza parziale, inferenza dinamica (asserzione, ritrattazione)
 - Descrizione di forme linguistiche e multi-modali
 - Lo studio dell'IA moderna non può prescindere dal calcolo simbolico
 - Inteso come strumento di analisi e di riferimento

Altre forme, altre discipline

- **Aspetti diversi, tecniche diverse**
 - Allo stesso tempo, lo stato attuale delle conoscenze in materia è organizzato per 'isole', ciascuna avente un proprio ambito di rilevanza
 - Oltre alle reti neurali, sistemi ibridi (reti probabilistiche, logiche *fuzzy* e non standard, algoritmi genetici, automi cellulari e molti altri) mostrano una notevole efficacia per classi di problemi particolari
- **Neuroscienze, sistemi biologici**
 - I modelli dell'IA influenzano e sono continuamente influenzati dai progressi degli studi nel campo della biologia, della neurofisiologia e delle scienze del comportamento e gli studi sulla percezione
- **Quadro di riferimento unitario**
 - Anche se un quadro complessivo di integrazione dei tanti aspetti dell'IA ancora non esiste
 - Vi sono numerosi tentativi di integrazione ed alcuni risultati cominciano ad emergere

I due corsi di Intelligenza Artificiale

- IA 1
 - Focalizzazione non esclusiva sui sistemi di calcolo simbolico
 - Logica classica proposizionale e del primo ordine (introduzione)
 - Sistemi a regole
 - E` prevalentemente orientato alla realizzazione di *agenti autonomi*
 - Strategie di rappresentazione delle informazioni
 - Metodi di ricerca, interpretazione dei dati sensoriali, pianificazione
- IA 2
 - Completamento dello studio della logica simbolica
 - Logiche del primo ordine e programmazione logica (cenni)
 - Logiche *fuzzy*, non standard ed applicazioni
 - Sistemi probabilistici, reti di Bayes
 - Automi cellulari, sistemi auto-organizzanti e adattativi
 - Algoritmi genetici e di calcolo evolutivo, per programmi e reti ibride
 - Realizzazione di *agenti autonomi* con tecniche avanzate