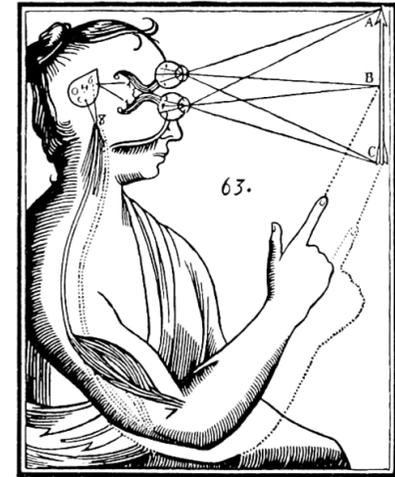


Intelligenza Artificiale I

Introduzione al corso

Calcolo e
rappresentazione

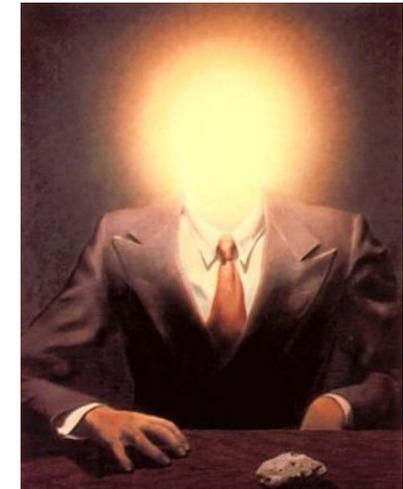
Marco Piastra



- 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 [...] 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.”



- 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.”

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.

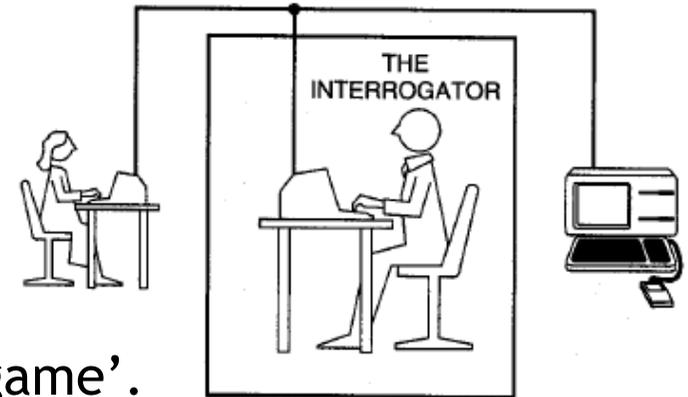
The object of the game for the interrogator is to determine which of the other two is the man and which is the woman.

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?’ ”



“Artificial Intelligence”, prima apparizione

- 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**.

features of intelligence...

- “*so precisely described*”

Si riferisce all'apprendimento o a qualsiasi altra caratteristica dell'intelligenza (umana)

- “that a machine can be made to *simulate* it”

Una descrizione estremamente precisa (e formale) è anche direttamente calcolabile?

- “a large part of human thought consists of manipulating *words*”

Quali parole (o simboli)?

Con quale significato?

- “according to *rules of reasoning* and *rules of conjecture*”

Regole di manipolazione delle parole (calcolo)

Regole separate per il ragionamento (certo) e le congetture (incerte)

Intelligenza e calcolo: *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)**

- Domande:

 - Deep Blue è intelligente?

 - Deep Blue *si comporta* in modo intelligente?

- 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>

Simulare il pensiero umano

- 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.”

(esempio di calcolo simbolico in matematica)

- Soluzione di equazioni algebriche

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

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

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

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

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

Una serie di passaggi: si applicano regole formali sulle espressioni

I simboli x , a e b *rappresentano* numeri reali

I simboli $+$, $-$, $/$, $*$, exp *rappresentano* operazioni sugli oggetti

Il simbolo $=$ *rappresenta* l'identità di due oggetti

L'equazione, nelle varie forme, esprime un legame di carattere generale

I passaggi effettuati sono mere operazioni sui simboli

La correttezza dei passaggi (regole formali) dipende dalle *rappresentazioni*

(esempio intuitivo di ragionamento simbolico)

- Fatti:
 1. Enrico è genitore di Paolo
 2. Silvia è genitrice di Giorgio
 3. Silvia è genitrice di Renato
 4. Enrico è 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 un genitore in comune sono parenti
 9. Chiunque abbia come parente un bipede senza piume è contento

(esempio intuitivo di ragionamento simbolico)

- 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 diverse? 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 genitore di Paolo
2. Silvia è genitrice di Giorgio
3. Silvia è genitrice di Renato
4. Enrico è 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 un genitore in comune sono parenti
9. Chiunque abbia come parente un bipede senza piume è contento

Calcolo simbolico

- Simboli

 - Sono (fisicamente) manipolabili da una macchina

 - Possono avere *relazioni* con altri simboli (sintassi: espressioni, formule)

- Elaborazione

 - La macchina utilizza *simboli* e *relazioni* ma opera solo sulla dimensione *fisica* (non ‘conosce’ il *significato** dei simboli)

- Processi

 - Il processo di ragionamento è una manipolazione dei simboli, secondo *regole* precise (calcolo)

 - L’implementazione della macchina che esegue il processo è irrilevante (*disembodiment* = il pensiero senza pensatore?)

- Calcolo simbolico *automatico*

 - Chi o che cosa esegue i processi?

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

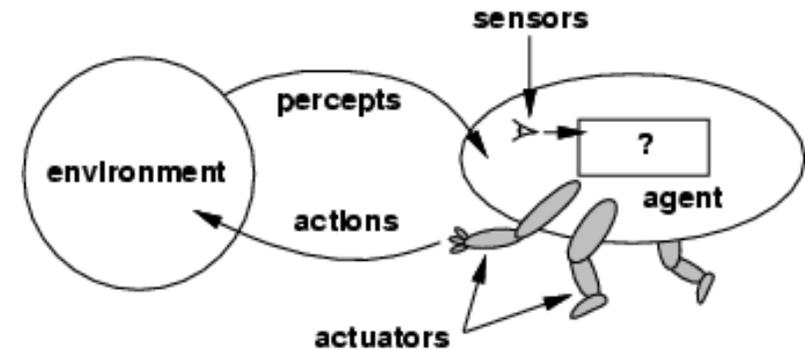
Comportamento, rappresentazione

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

“[...] 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.”

Comportamento razionale



- Agenti

Il sistema (la macchina) è un **agente** che interagisce con un ambiente esterno

Elaborazione: un ciclo ripetuto (in sequenza o in parallelo) $f: P^* \rightarrow A$

Percezione (attraverso i sensori, P^*)

Valutazione (processi interni $f: P^* \rightarrow A$)

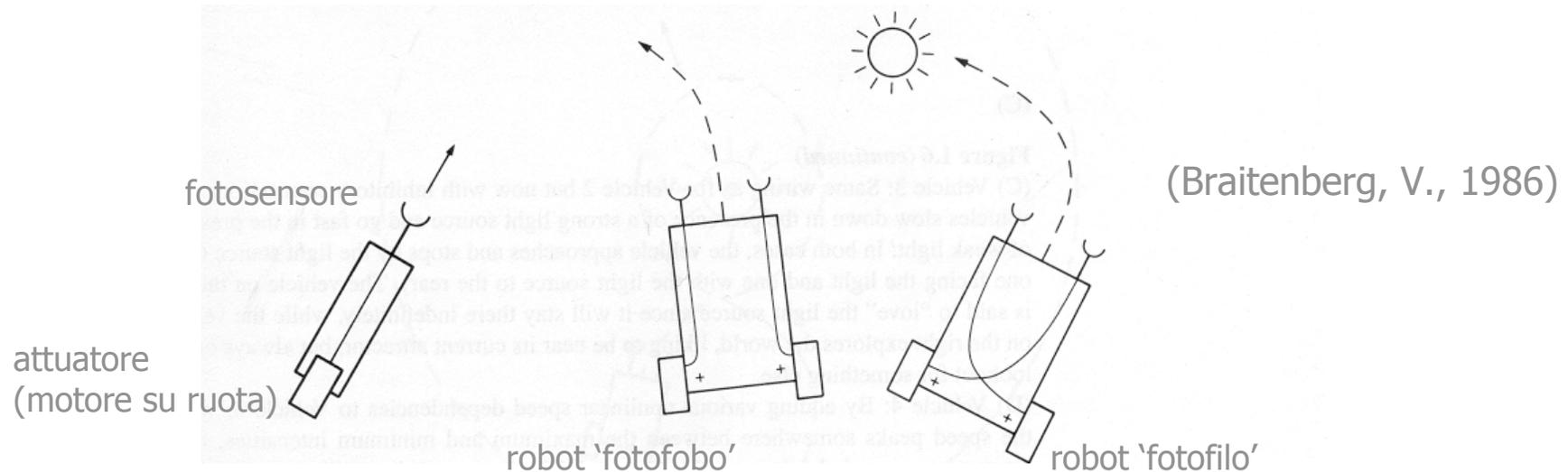
Azione (attraverso gli attuatori, A)

- Come si determina la razionalità dell'agente?

Criteri possibili:

- La fase di valutazione avviene in modo razionale
(ad esempio l'agente adotta una rappresentazione interna simbolica)
- “Fa’ la cosa giusta”
(Si considera il comportamento osservabile a prescindere dai processi interni)
- “Agisce come un essere umano”
(Si confronta il comportamento osservabile con un campione)

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

Gli elefanti sanno giocare a scacchi?

(Brooks, R., *Elephants Don't Play Chess*, 1990)

- Critica dell'ipotesi del sistema di calcolo 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

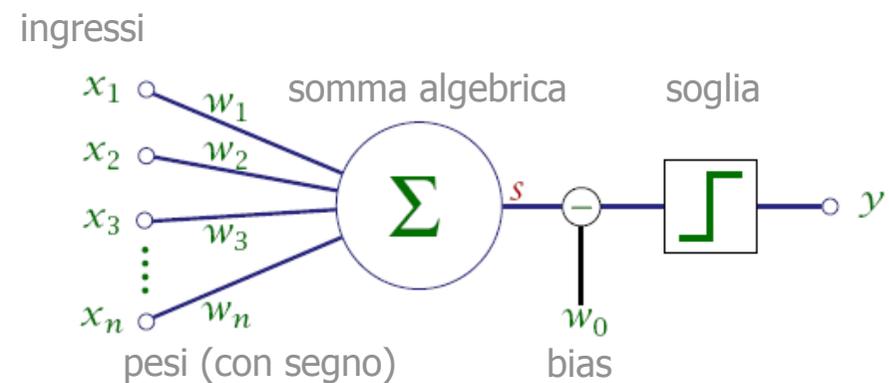
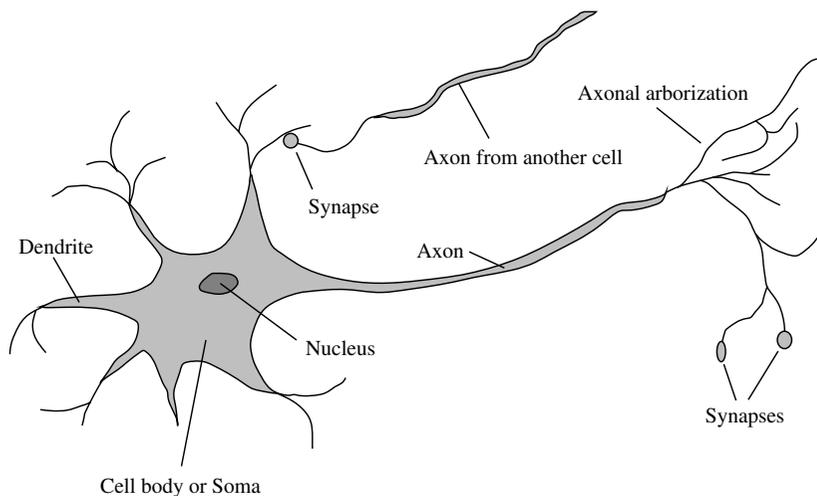
Connessionismo

“Occorre considerare anche il *cervello*, non solo la mente”
 un’ipotesi opposta al *disembodiment* tipico del ragionamento simbolico

- **Neuroni artificiali derivati 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)

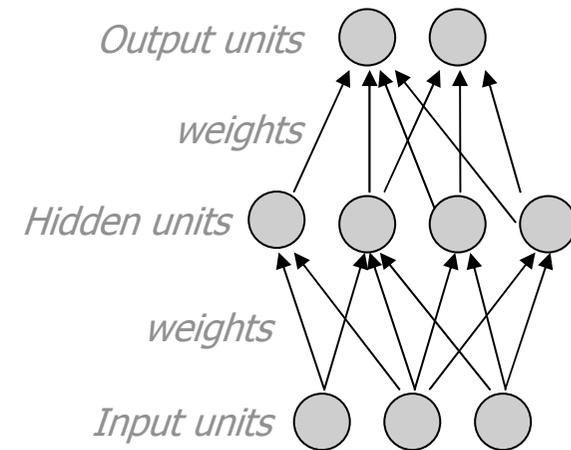
Non esiste una rappresentazione interna immediatamente riconoscibile

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

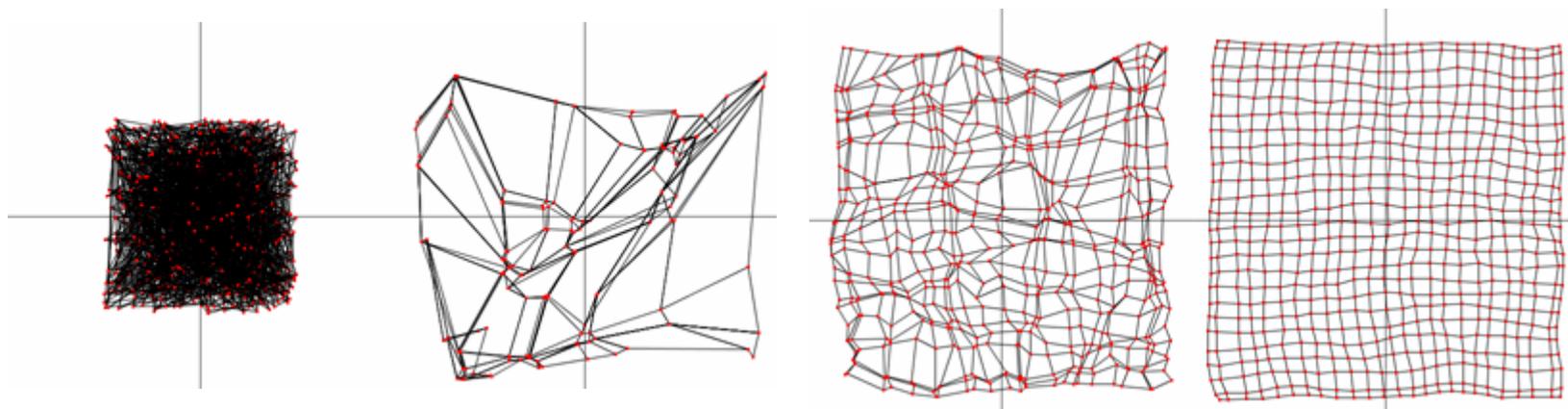
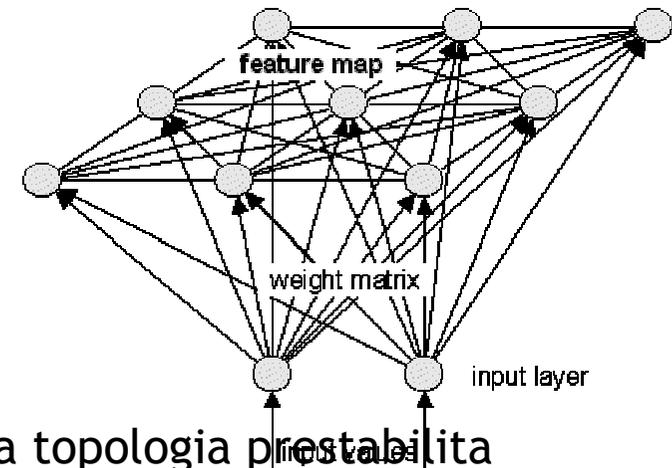
Livello di mappa, con unità organizzate secondo una topologia prestabilita

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'



Comportamento e rappresentazione 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)

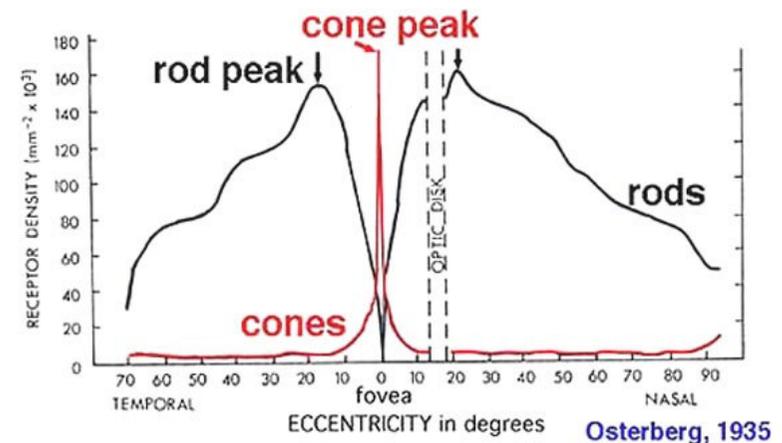
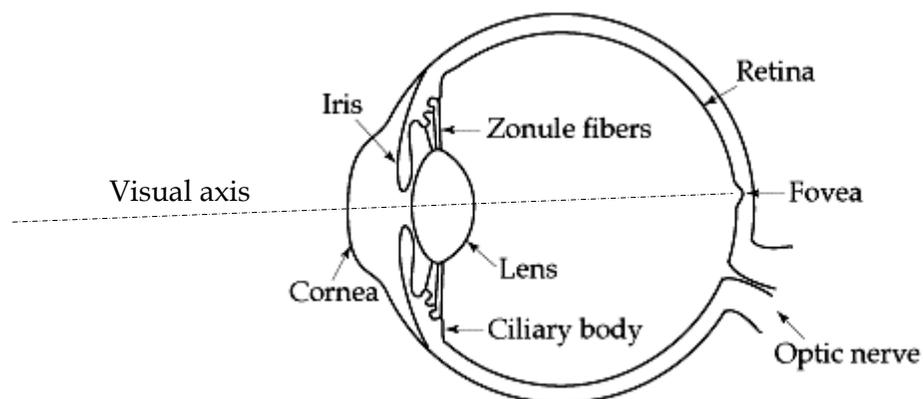
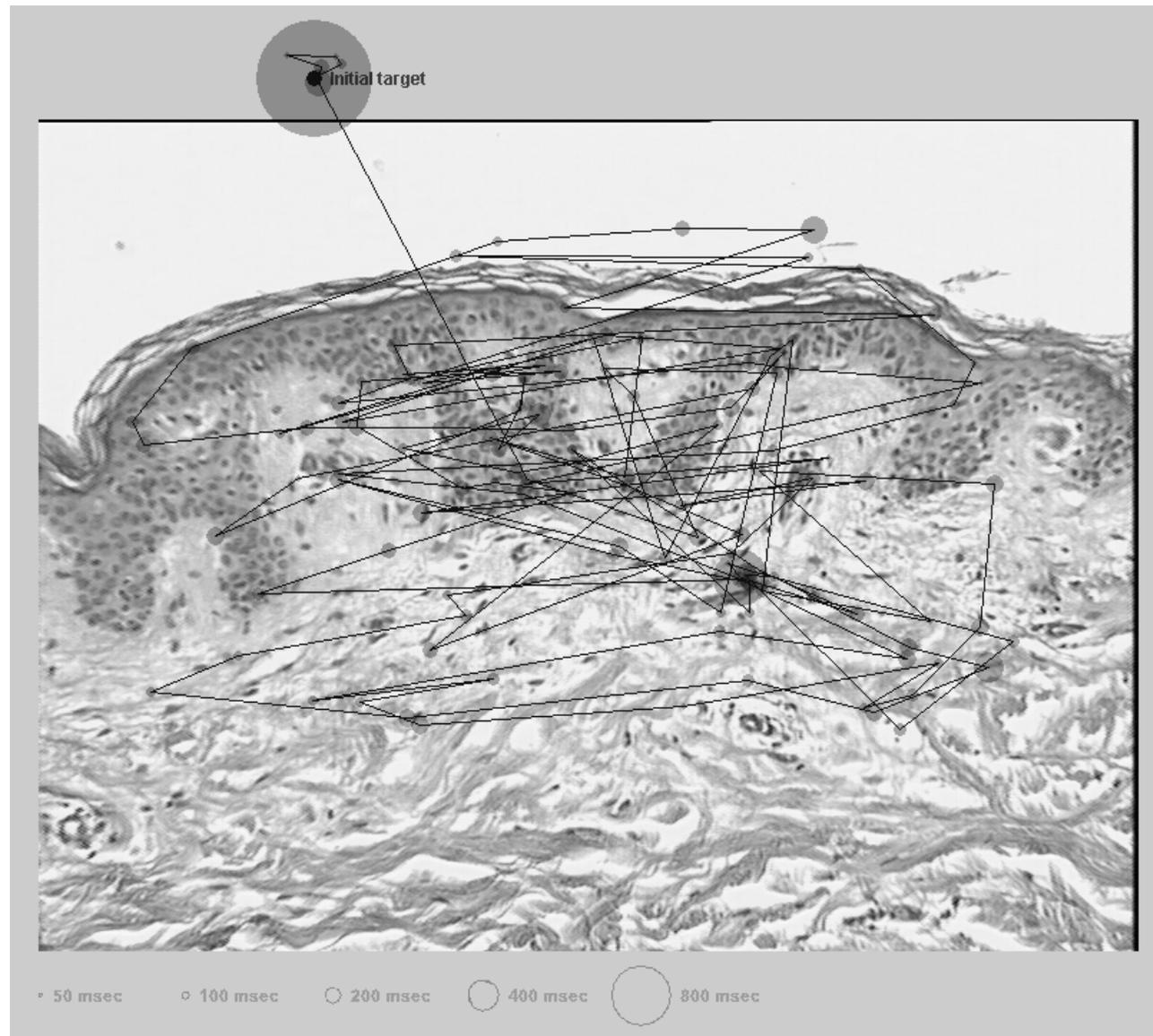


Immagine al microscopio di una sezione di cute opportunamente trattata

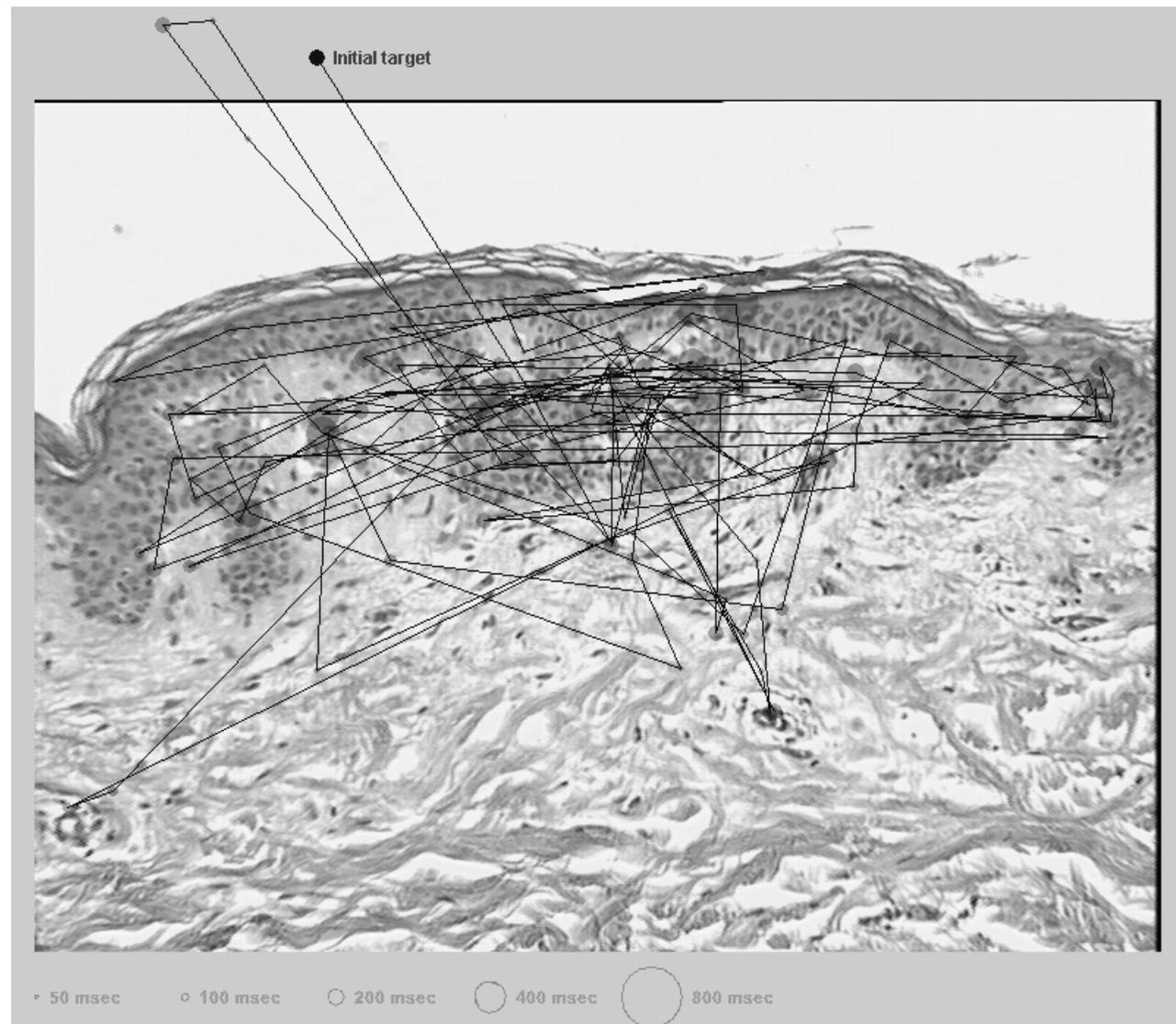
Lo scopo dell'osservazione è una *diagnosi*

I soggetti vedono le immagini per la prima volta

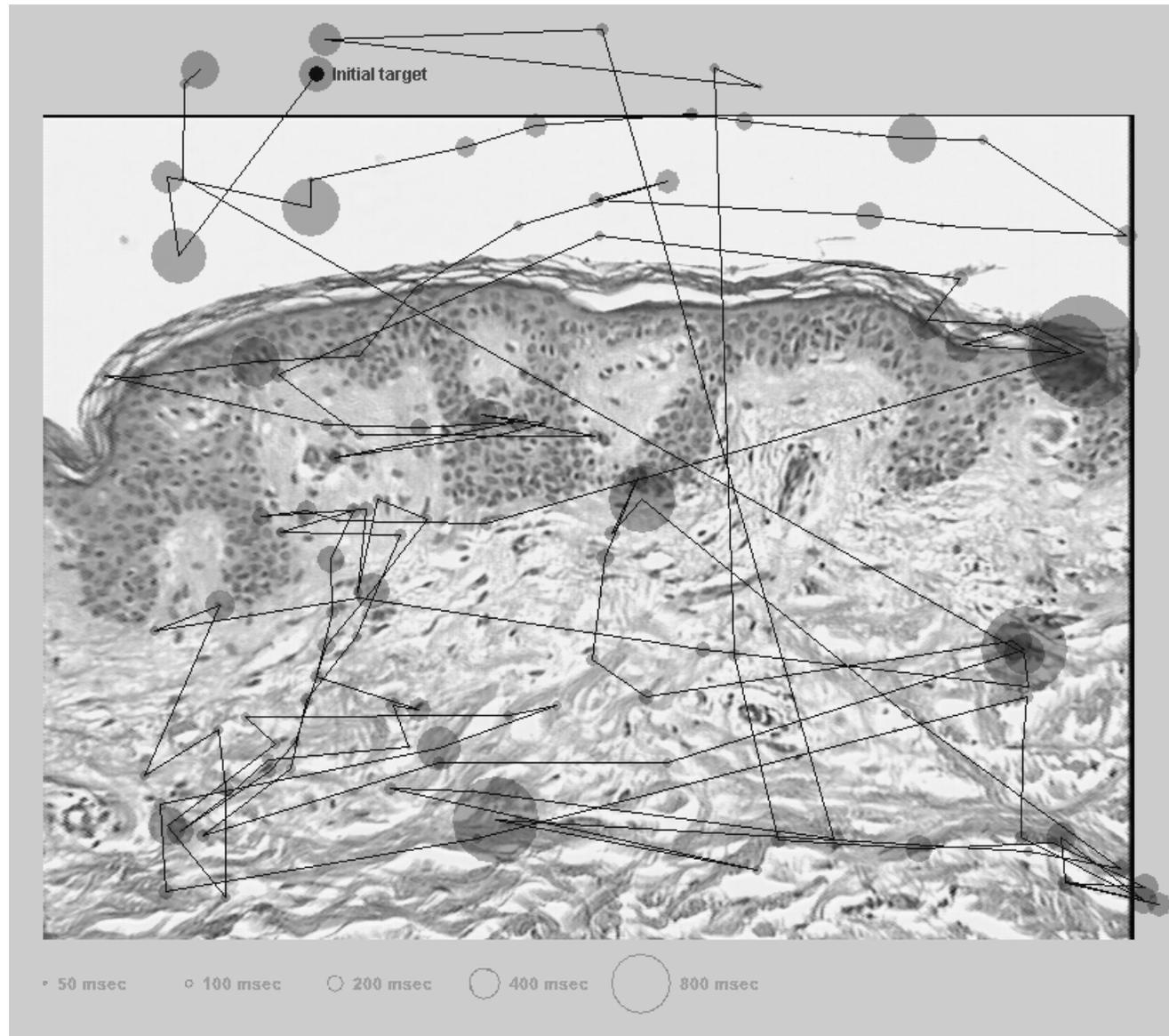
Lo *scanpath* tracciato è di un giovane medico (durata: 45 sec)



Lo *scanpath*
di un medico
molto esperto

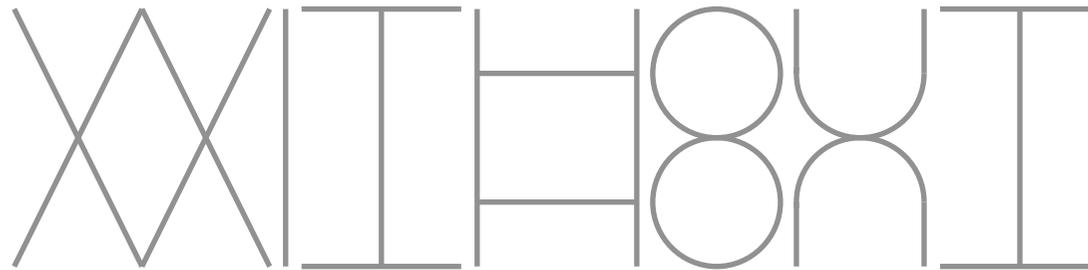


Lo *scanpath* prodotto da un ingegnere (non conosce l'argomento)



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

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:

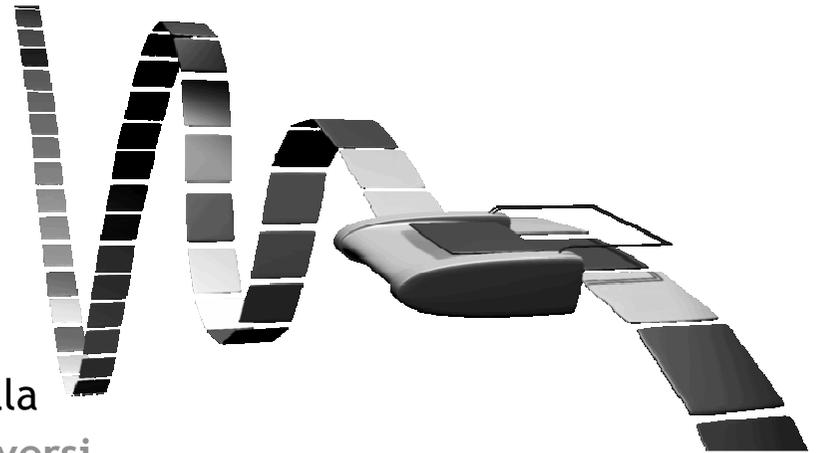
“X ama Y”

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)

Macchina di Turing



- Un modello astratto

- Un nastro con celle elementari, un simbolo in ogni cella

- Una testina di lettura e scrittura delle celle, può muoversi

- Uno stato della macchina (un simbolo)

- Una CPU, governata da una tavola di transizione

- { <stato attuale, lettura cella> → <azione, nuovo stato> }

- (dove azione può essere scrittura cella o movimento)

- Generalità

- Descrive la potenza di calcolo di un computer qualsiasi

- Qualunque *funzione* sia calcolabile da un computer, è calcolabile da una macchina di Turing

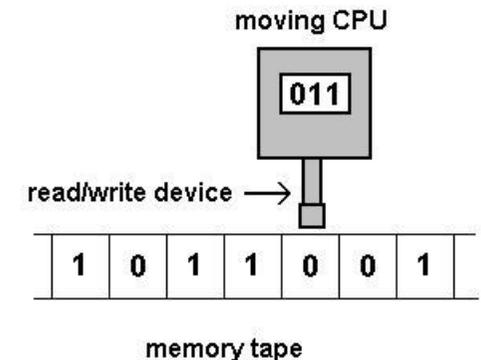
- L'unica differenza idealizzazione è il nastro illimitato

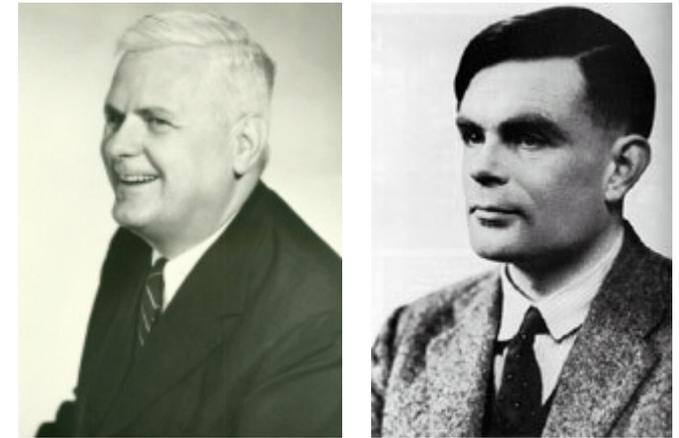
- Macchina di Turing Universale (*Universal Turing Machine*)

- Versione estesa della macchina di Turing

- La tavola di transizione viene caricata dal nastro, all'inizio

- Può emulare qualsiasi macchina di Turing, data la tavola di transizione





Tesi di Church-Turing

Non esiste un'unica, concisa formulazione originale:
si tratta di un concetto espresso in più passaggi, da integrare con risultati teorici

- Una possibile formulazione (Wikipedia)

“Every 'function which would naturally be regarded as computable' can be computed by a Turing machine.”

La vaghezza della formulazione ha dato luogo a diverse interpretazioni,
una molto comune (e non del tutto esatta) può essere espressa come in (Wikipedia):

“Every 'function that could be physically computed' can be computed by a Turing machine.”

Searle: “... At present, obviously, the metaphor is the digital computer.”

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)

- **Significato (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* (= in cui le variabili assumono 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

Logica come calcolo simbolico

▪ Rappresentazione, calcolo

Una disciplina antichissima (da Aristotele in poi, ~320 a.c.)

Si studia solo la caratterizzazione moderna

Linguaggio

Simboli elementari che si compongono in *formule*, secondo una sintassi rigorosa

Semantica

Ai simboli elementari e quindi alle formule è attribuito un significato univoco

Astrazione formale (per generalizzazione)

Il significato delle formule coincide con una pluralità di possibili stati delle cose, o *mondi possibili* (una formula può essere applicabile a contesti diversi)

Regole di inferenza

Stabiliscono in modo preciso quali formule possono essere *derivate* da quali altre
“Mai una conseguenza falsa da una premessa vera!”

Ragionamento come processo di derivazione

Si parte da un'insieme di formule, si applicano iterativamente le regole
(*se va bene*, si arriva al risultato)

Logica come calcolo simbolico *automatico*

■ Processi di calcolo

Regole diverse, stessi risultati

Il calcolo simbolico automatico richiede regole di inferenza speciali

Il risultato dell'applicazione delle diverse regole deve essere lo stesso (almeno in teoria)

Risolvibilità dei problemi

Non tutti i problemi ammettono una metodo di soluzione generale (p.es. eq. diofantine)

Non tutti i metodi di soluzione generale sono applicabili da una macchina (in tempo finito)

Non tutti i metodi applicabili in tempo finito sono adeguatamente efficienti

Tecniche di calcolo

Come si applicano le regole di inferenza?

Come si può ottimizzare l'esecuzione del processo?

Tecniche e problemi (tipi di ragionamento)

Alcune tecniche sono più adatte di altre a problemi di un tipo specifico

I due corsi di Intelligenza Artificiale

- IA 1: Calcolo e Rappresentazione

 - Logica di base e calcolo automatico

 - Logica classica proposizionale e del primo ordine

 - Sistemi a regole

 - Realizzazione di *agenti autonomi*

 - Strategie di rappresentazione delle conoscenze

 - Metodi di ricerca, interpretazione dei dati sensoriali, pianificazione

- IA 2: dal Calcolo Simbolico al *Natural Computing*

 - Logica avanzata

 - Logiche del primo ordine e programmazione logica (cenni)

 - Logiche *fuzzy*, non standard ed applicazioni

 - Sistemi probabilistici, reti di Bayes, apprendimento

 - Automati cellulari, sistemi auto-organizzanti e adattativi

 - Algoritmi genetici e di calcolo evolutivo

 - Realizzazione di *agenti autonomi* con tecniche avanzate

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.”