

# Intelligenza Artificiale

2024-2025

L'intelligenza artificiale è una disciplina ampia con diverse tecniche e diversi obiettivi.

*Ragionamento*: risolvere problemi nuovi sfruttando conoscenze. Intelligenza  $\neq$  personalità.

L'intelligenza artificiale è una branca dell'informatica che progetta sistemi con capacità generalmente considerate come riservate all'uomo. È scienza (studio teorico, anche comparativo con l'intelligenza umana) e ingegneria (produce applicazioni concrete utili).

Le intelligenze artificiali ottenute tramite *machine learning* sono diventate così complesse da essere studiate come *black-box*, tecnica generalmente riservata ai fenomeni naturali.

Alan Turing dimostrò che è possibile costruire una macchina per ogni algoritmo definito. Ideò anche il test di Turing per intelligenze artificiali nel 1950. Neumann, oltre a definire la famosa architettura, fu il primo a teorizzare programmi autoreplicanti e capaci di evolversi come gli organismi naturali.

La *cibernetica* confronta i sistemi artificiali agli animali.

Nel 1943 McCulloch e Pitts creano il primo modello matematico di *neuroni artificiali*. Nel 1949 Hebb teorizza l'apprendimento autonomo. Solo di recente questi concetti hanno iniziato ad avere utilità pratica grazie all'aumento di disponibilità di dati e potenza di calcolo.

Il termine *intelligenza artificiale* viene definito nel 1956 a Dartmouth da McCarthy, Minsky, Rochester, Shannon ed Elwood. Nascono due paradigmi:

- paradigma di simulazione
- paradigma prestazionale o di emulazione

I primi programmi AI creati dal team soffrono di problemi di scalabilità. Viene per la prima volta riconosciuto il problema della necessità di conoscenza.

Dendral (1969) è il primo esempio di *sistema esperto* per la classificazione di sostanze chimiche basato sui dati spettrometrici. I sistemi esperti sono agenti basati sulle conoscenze e sono costituiti da regole codificate manualmente da esperti del settore.

I sistemi esperti sono esempi di *intelligenza artificiale simbolica*. Sono costruiti con oggetti astratti (*simboli*) e regole logiche.

Un agente opera in maniera autonoma, percependo l'ambiente, adattandosi ai cambiamenti e perseguendo obiettivi. La definizione di ambiente è flessibile e si può riferire anche ad un ambiente completamente virtuale, come l'interno di un computer.

Un esempio di regola è la seguente:

Un professore può lavorare per un solo ateneo per volta.

In termini logici:

$$\text{Professore}(x) \wedge \text{LavoraPer}(x, y) \wedge \text{LavoraPer}(x, z) \wedge x \neq z \rightarrow \perp$$

Mettiamo di avere la seguente conoscenza pregressa (*stato*):

$$KB = \{\text{Professore}(P), \text{LavoraPer}(P, U1)\}$$

e di ricevere una nuova informazione:

$$I = \{\text{LavoraPer}(P, U2)\}$$

Possibili approcci:

- cambiare  $KB$  per rispettare la regola in funzione di  $I$ , due mutazioni:
  - $P$  non è più un professore
  - $P$  si è licenziato da  $U1$  per lavorare presso  $U2$
- ignorare  $I$  a favore della conoscenza pregressa in  $KB$

Se non è definito in modo chiaro il percorso da prendere, il sistema si blocca. Gli umani, invece, sono adattati a lavorare con stati inconsistenti e incerti senza bloccarsi perché hanno una conoscenza probabilistica della realtà.

Il ragionamento serve per mantenere lo stato consistente dopo nuove percezioni, e per decidere azioni verso un obiettivo.

Gli agenti hanno un ambito limitato e specifico. Esempi storici includono MYCIN per la diagnostica del sangue, con metriche probabilistiche di incertezza, SHRDLU per il linguaggio naturale, o il sistema LUNAR per le rocce lunari.

Il 1979 vede la nascita delle scienze cognitive. Nel 1981 il Giappone lancia il piano di investimenti Fifth Generation. Negli anni '80 Xerox, Texas Instruments e Symbolics realizzano workstation per il linguaggio LISP. Nel 1986 quattro diversi gruppi indipendenti riscoprono l'algoritmo di retropropagazione. Nel 1988 viene inventato l'apprendimento probabilistico. Gli anni 2000 vedono la nascita del *big data*, grandi dataset raccolti tramite Internet. Il 2011 marca l'apertura dell'epoca del *deep learning*.

L'AI tradizionale (simbolica) ha una visione algoritmica e deduttiva sia dei sistemi artificiali che della mente umana, mentre l'approccio moderno è induttivo e basato su processi incrementali. Negli anni recenti si afferma l'idea che i sistemi migliori nascano dall'unione di sistemi simboli e sistemi ad apprendimento automatico, con un'unione di informazione e semantica.

La questione di fondo rimane intoccata: come rappresentare la conoscenza interna? Quali processi razionali rappresentare?

Si sviluppano metodi e teorie: rappresentazione della conoscenza, ragionamento automatico, pianificazione... Al contempo nascono soluzioni pratiche: computer vision, programmazione in linguaggio naturale, sistemi esperti e modelli.